Evidence for a new boson in the search for the Standard Model Higgs particle decaying to four leptons at CMS

Dott. Marco Meneghelli

Advisors:

Prof. F.L. Navarria Dott. F.R. Cavallo

Dott. P. Giacomelli Dott. S. Braibant

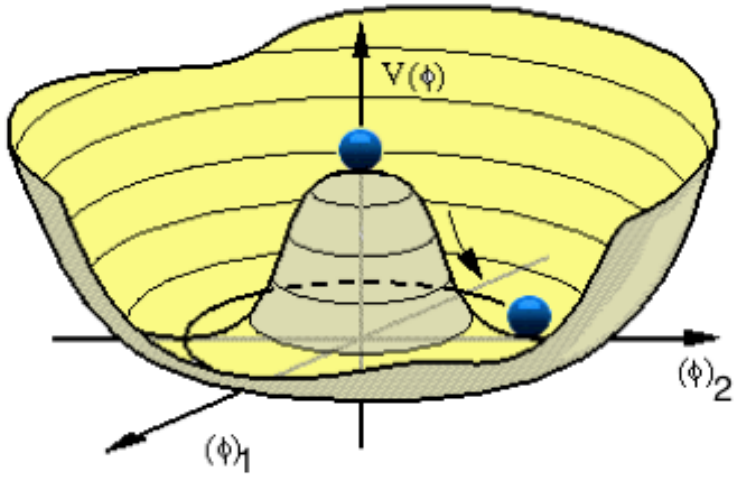
The Standard Model (SM) and the Higgs boson

- **The SM of elementary interactions is a QFT built on symmetry principles**:

$$SU_c(3) \times SU_L(2) \times U_Y(1)$$

- It describes matter as composed by **fermion fields that interact exchanging vector bosons**
- An **elementary scalar field** is present in the theory
- The **electroweak spontaneous symmetry breaking mechanism** gives rise to the mass of the bosons W and Z and predicts the **existence of a scalar boson: the Higgs**
- The scalar boson couples also to fermions, via Yukawa-like interactions, introducing mass-terms in the theory Lagrangian
- **The mass of the Higgs boson is a free parameter of the theory**

$$V(\phi^\dagger\phi) = \lambda(\phi^\dagger\phi)^2 + \mu^2\phi^\dagger\phi$$



$$SU_L(2) \times U_Y(1) \rightarrow U_{em}(1)$$

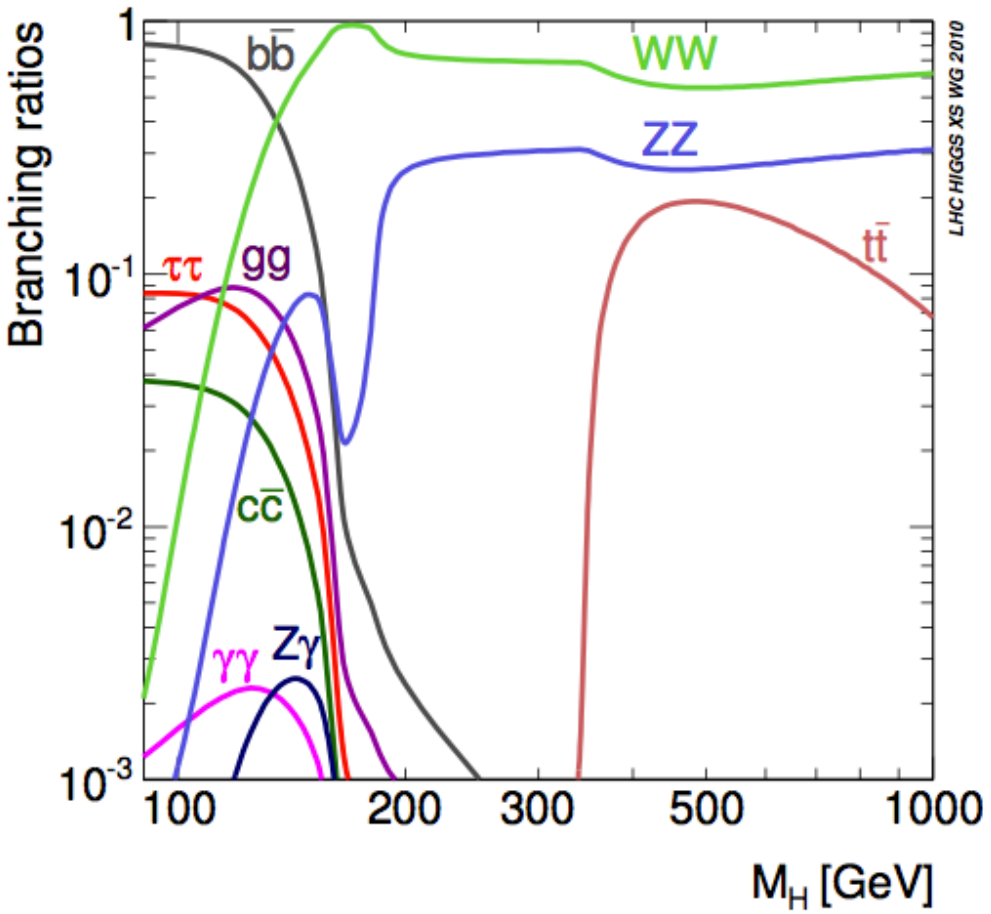
Search for the Higgs boson at the LHC

- **The LHC** is a proton-proton collider:

  CM energy: 7 - 8 TeV, Luminosity: O(10^{33}-10^{34}) $cm^{-2}s^{-1}$

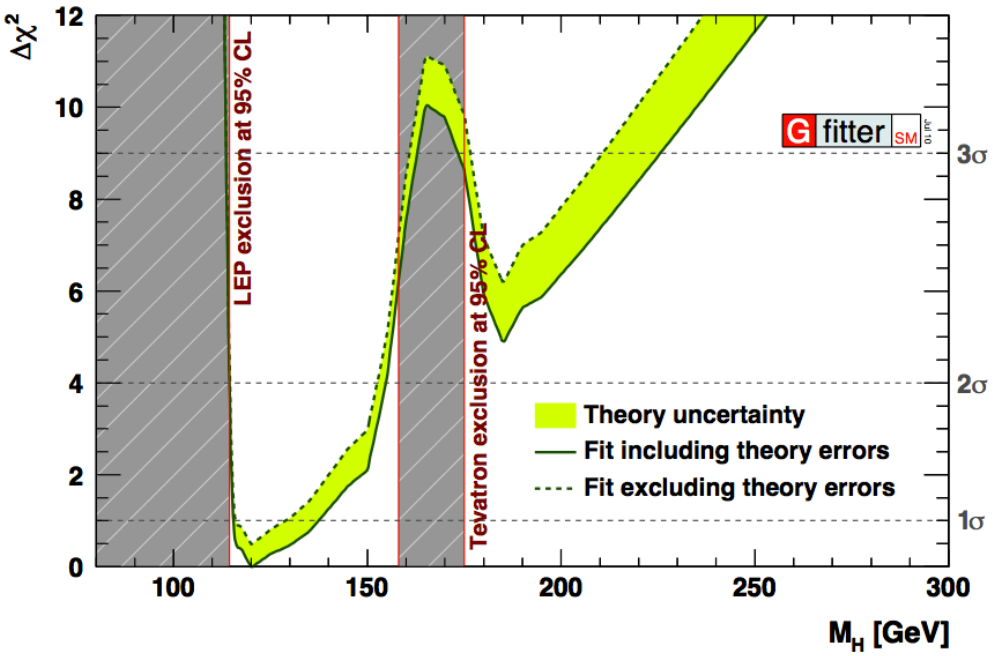
- Main difficulty is to discriminate signal processes against the background
  - very different final states

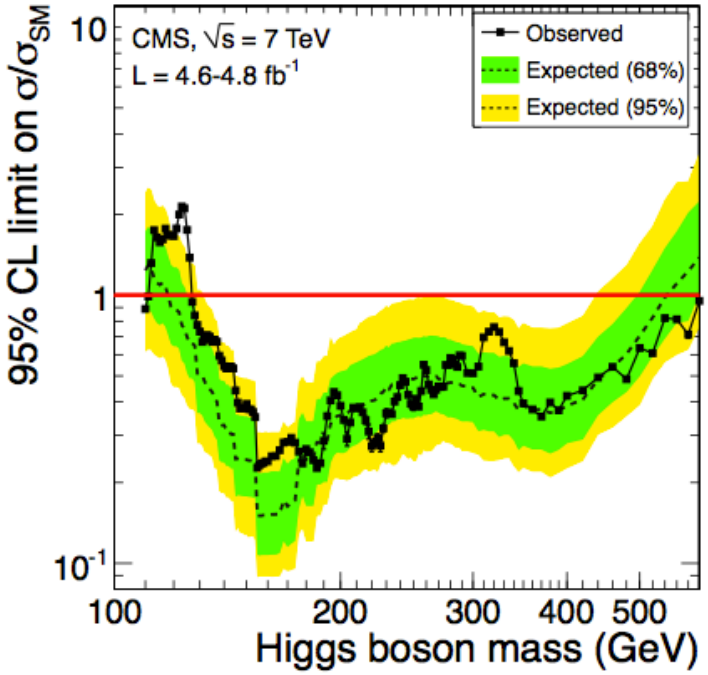
- Higgs production
  - Gluon fusion: gg→H
  - V B F: VV→H
  - Associate prod: HV
  - Associate top: Htt
- **Best search channels**:
  - **H→ZZ→4l**
  - **H→γγ (low mass)**
- Other search channels:
  - H→bb (low mass)
  - H→WW→2l2v
  - H→ττ (low mass)

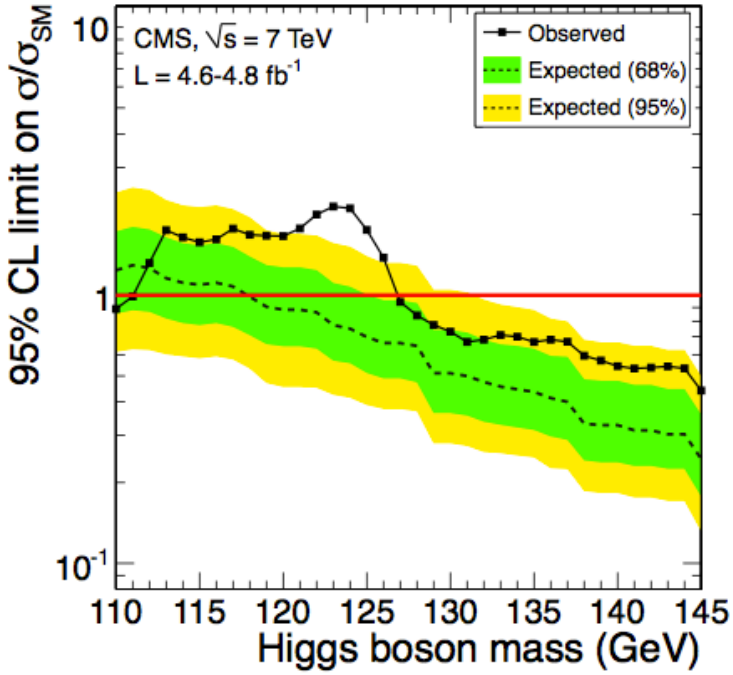


Up-to-2012 scenario

- up to 2011:
  - LEP and Tevatron **direct exclusions** (95% CL)
  - **Best fit** to EW/SM data suggests a light Higgs
- **2011:**
  - CMS excludes m_H in [**127**,600] GeV
  - ATLAS excludes m_H in [111.4,116.6] U [119.4,122.1] U [**129.2**,541] GeV

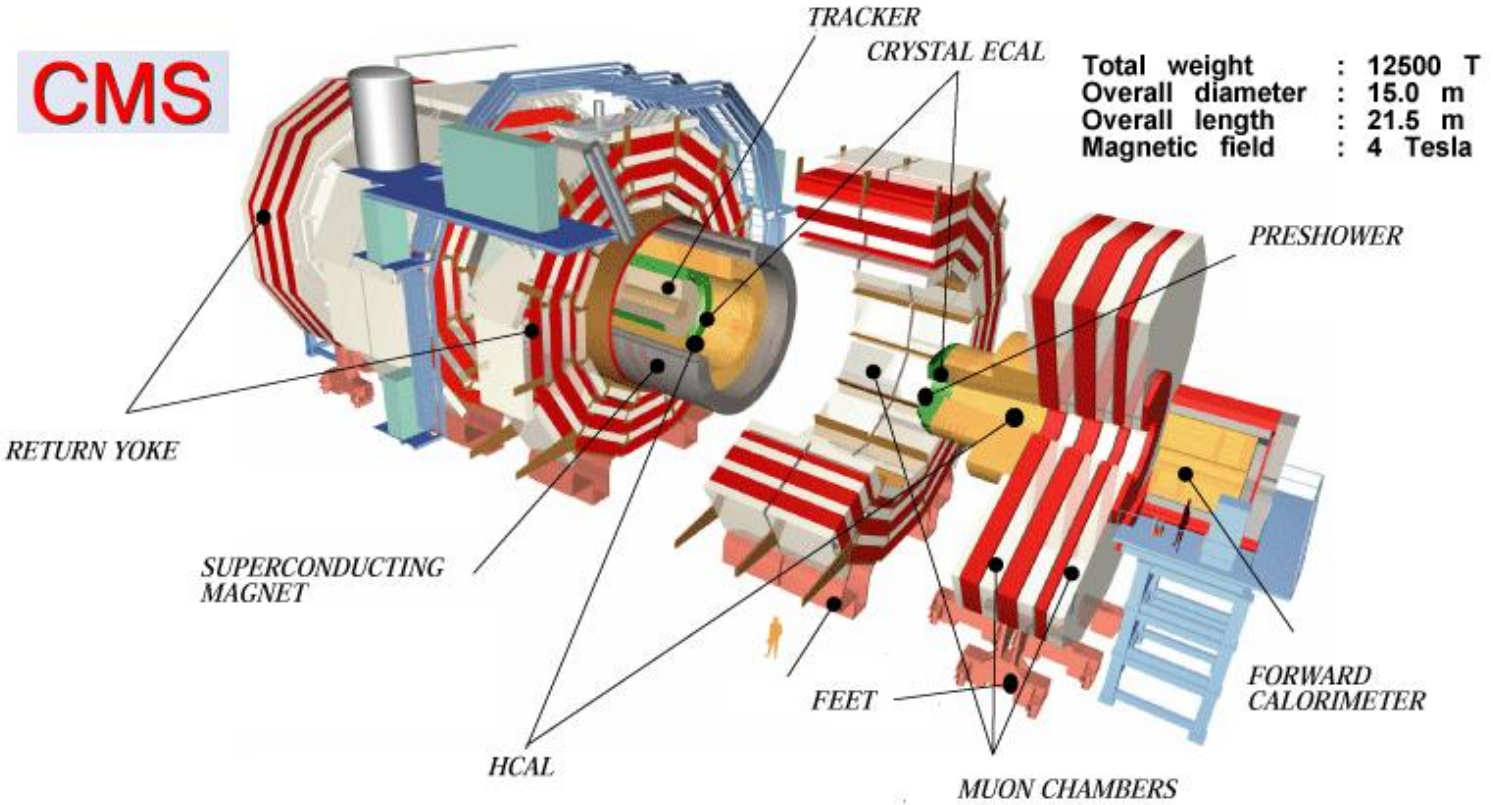






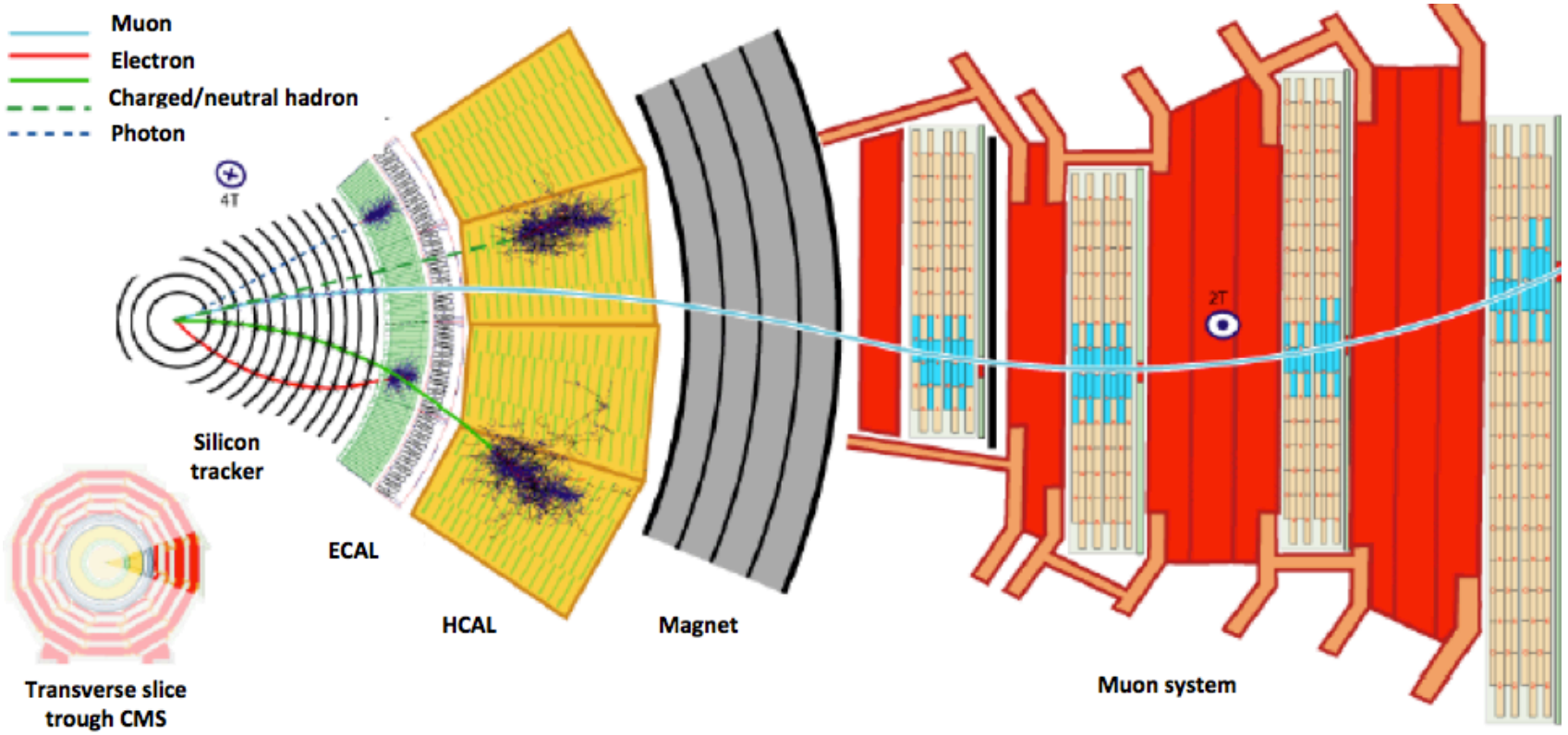
The Compact Muon Solenoid (CMS) experiment

- **Studies the p-p collisions at high energy**
- Compact cylindrical structure, built around the solenoid (3.8 T)
- Central silicon tracker
- Electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL)
- Redundant muon system

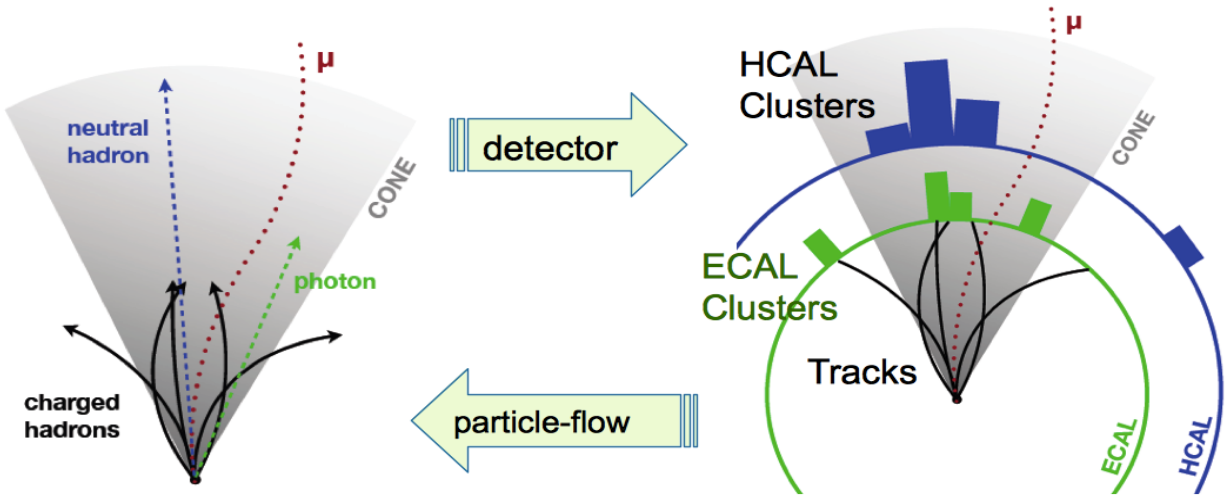


Physical object reconstruction at CMS

- The particles are identified from the different behavior in the detector
- **Particle Flow algorithm** yields a global description of the event



Lepton reconstruction and identification at CMS



- Global and tracker **muons**, reconstructed matching **tracker & muon system** information (p_T>5, |η|<2.4)
- Particle Flow Muons
  - requirements on track components
  - energy deposits in the calorimeters

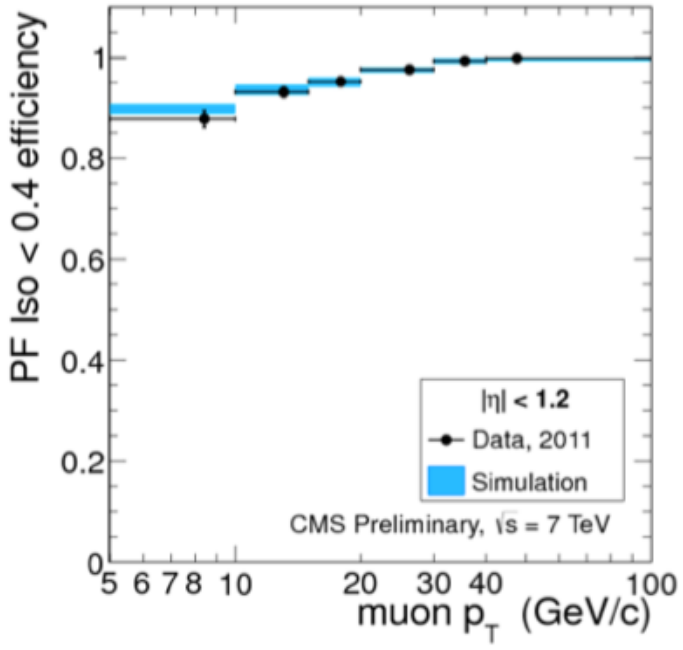
Efficiency ~ 98%

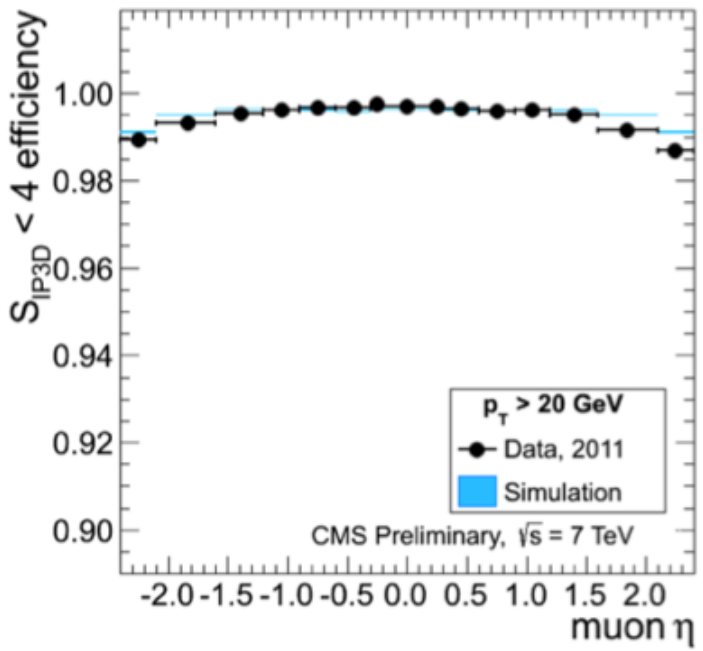
- Gaussian Sum Filter **electrons**, reconstructed matching **tracker & ECAL** information (p_T>7, |η|<2.5)
- multivariate BDT identification:
  - bremsstrahlung
  - matching tracker-ECAL
  - shower shape

Efficiency ~60 → 85 % (p_T <~ 10 GeV)
~90→ 95 % (p_T >~ 20 GeV)

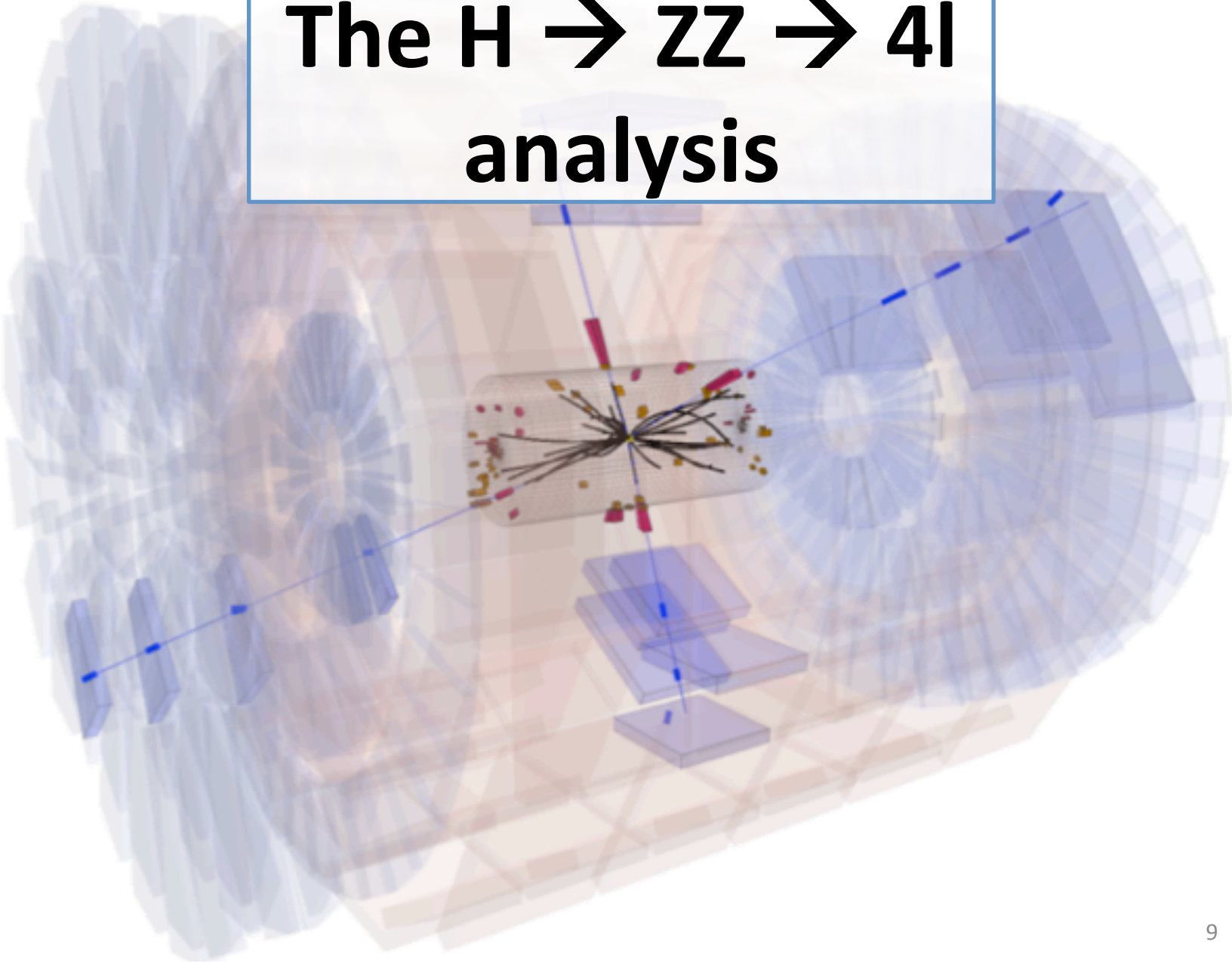
Prompt lepton selection

- The leptons used in the analysis are required to be **isolated**, i.e. not contained in a jet
  - Lepton isolation is evaluated summing the p_T of tracks or energy deposits in a cone around the track of each lepton, and normalizing to the lepton p_T (**Iso**)
- The leptons are required to come from a **common primary vertex** to discriminate against secondary leptons
  - **Significance of impact parameter**: $|SIP_{3D}| = |IP_{3D}/s_{IP}| < 4$

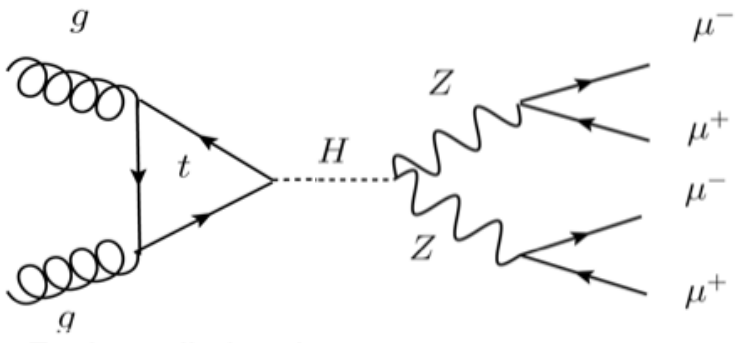


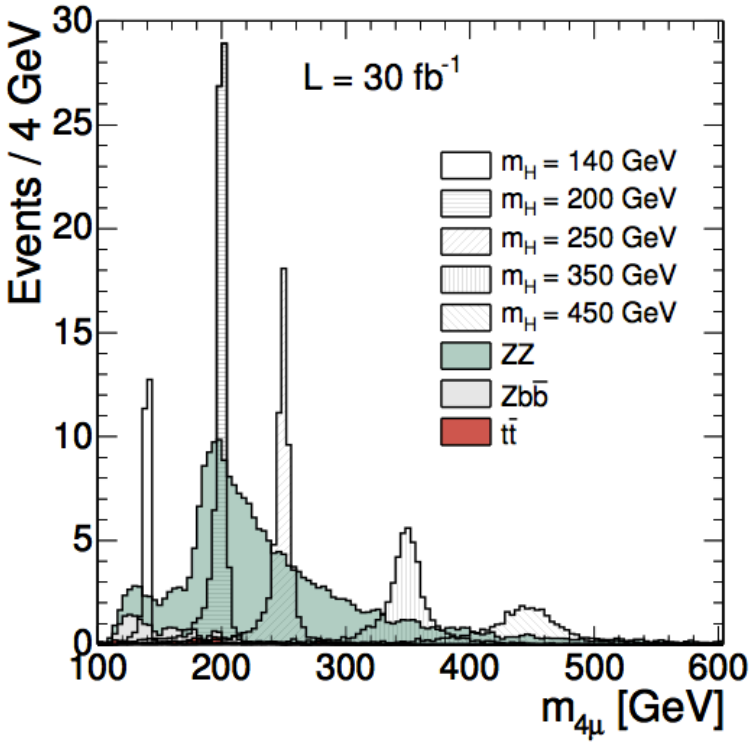


The H → ZZ → 4l analysis



The 4-lepton channel



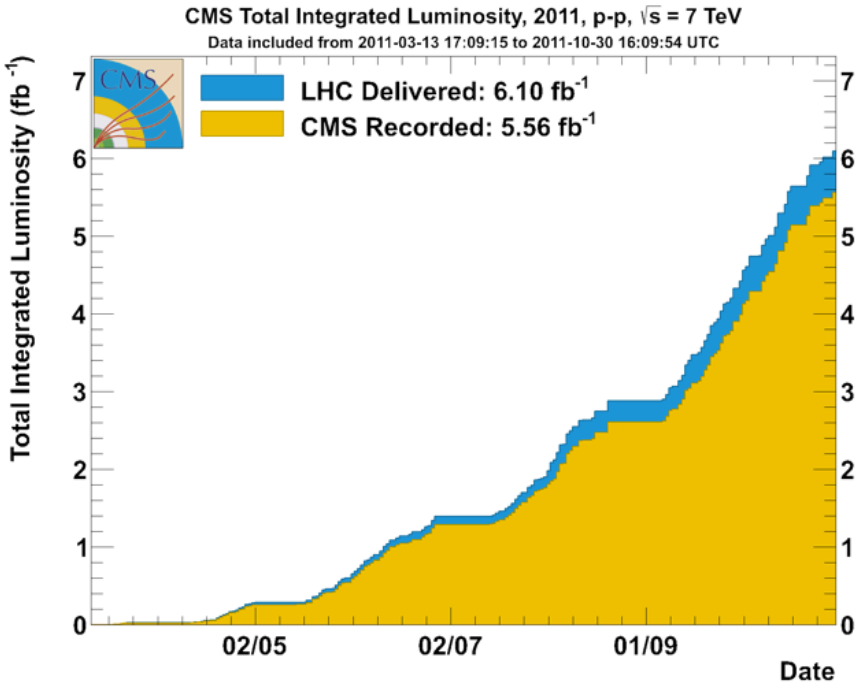


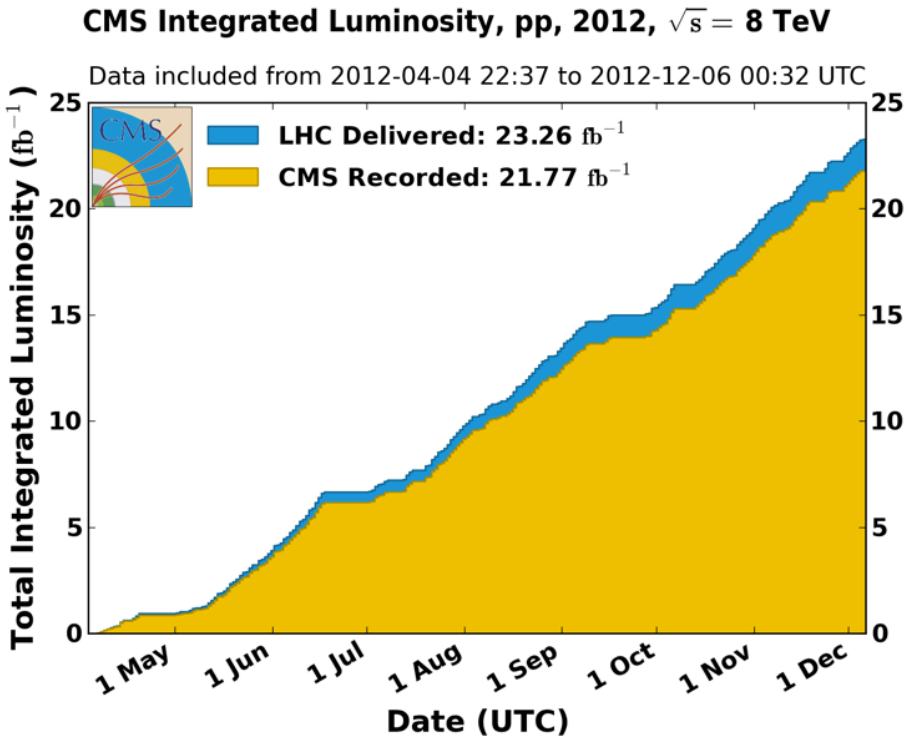
- Decay channel: $H \rightarrow ZZ^* \rightarrow l_1^+ l_1^- l_2^+ l_2^-$ $\quad l_i = \mu, e$
- **Experimental signature:**

  4 isolated leptons coming from the Primary Interaction Vertex
- **Resonance peak** in the 4-lepton mass distribution with excellent resolution, O(GeV)
- Ratio **signal:background** locally of the order **1:1**
  - Irreducible background: the **ZZ continuum**
  - Reducible backgrounds: **Z+jets, tt, WZ, QCD**
- Good sensitivity in a wide mass range $115 < m_H < 700$ GeV

Data and simulations

- Public data, collected in 2011 and 2012:
  - **5 fb^{-1}** with protons colliding **at 7 TeV**
  - **12 fb^{-1}** with protons colliding **at 8 TeV**
- MC simulations of signal and background processes using Pythia, Powheg, Madgraph and a full detector simulation with Geant4





$Z \rightarrow l^+l^-$: event selection

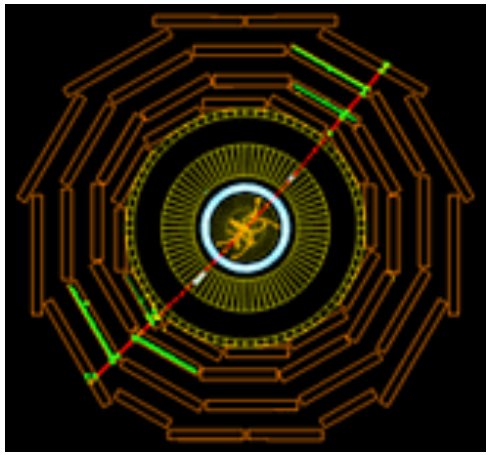
- Before going to the 4-lepton analysis it is **useful to study the single Z decay to two leptons** → high statistics → highlights the systematics
- **Event selection:**

Trigger
double lepton trigger paths required

Photon selection

FSR recovery

Lepton selection
ID criteria
Isolation and SIP

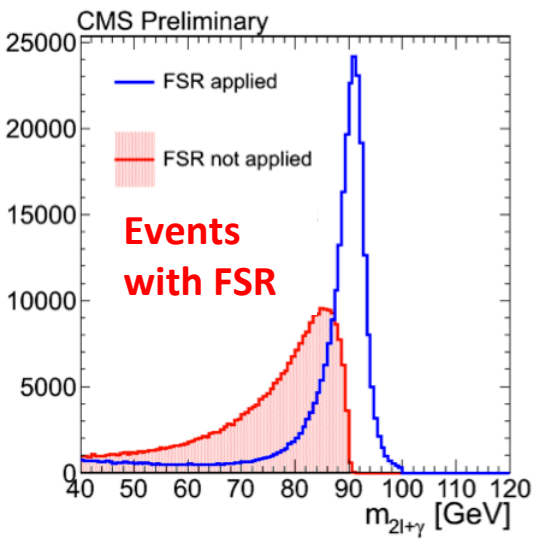


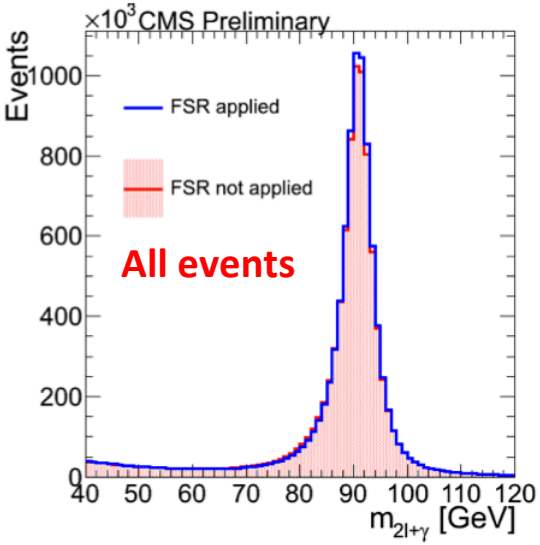
Build Z candidate
lepton pair with invariant mass closest to 91.19 GeV

Final State Radiation (FSR) recovery

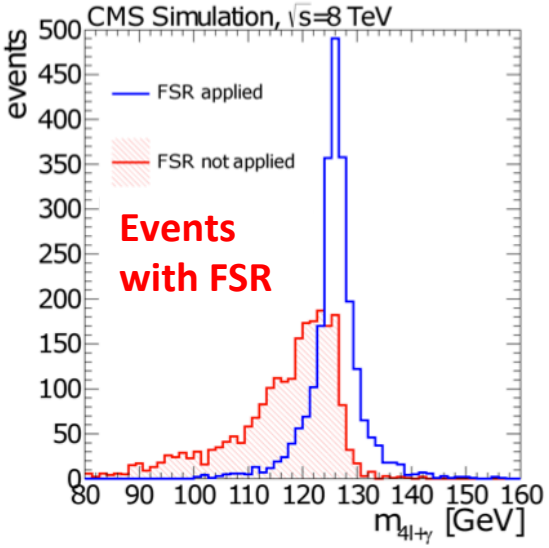
Associating FSR photons to Z candidates → gain ~2-3% of signal efficiency

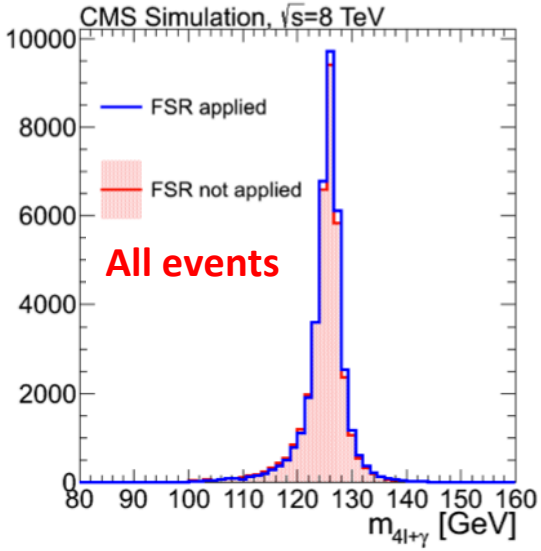
Single Z
(2011 data)





Higgs 126 GeV
(MC)

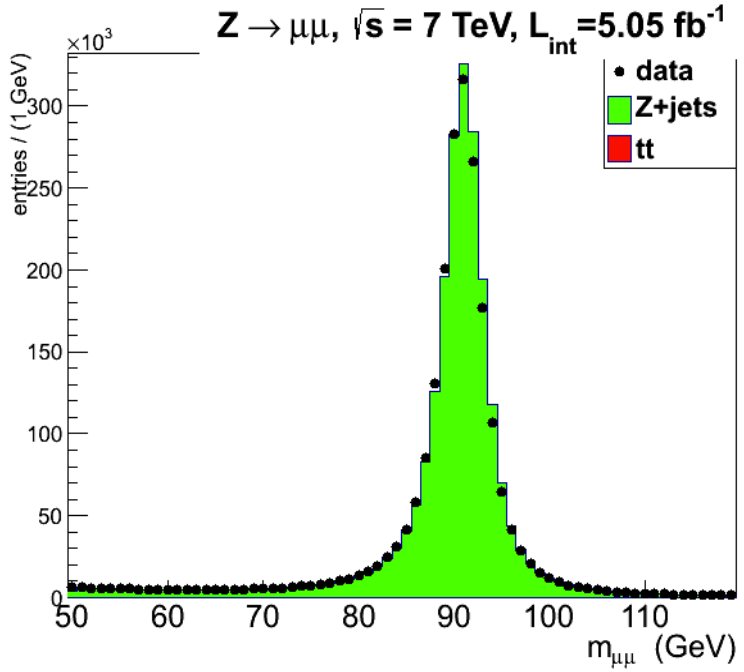


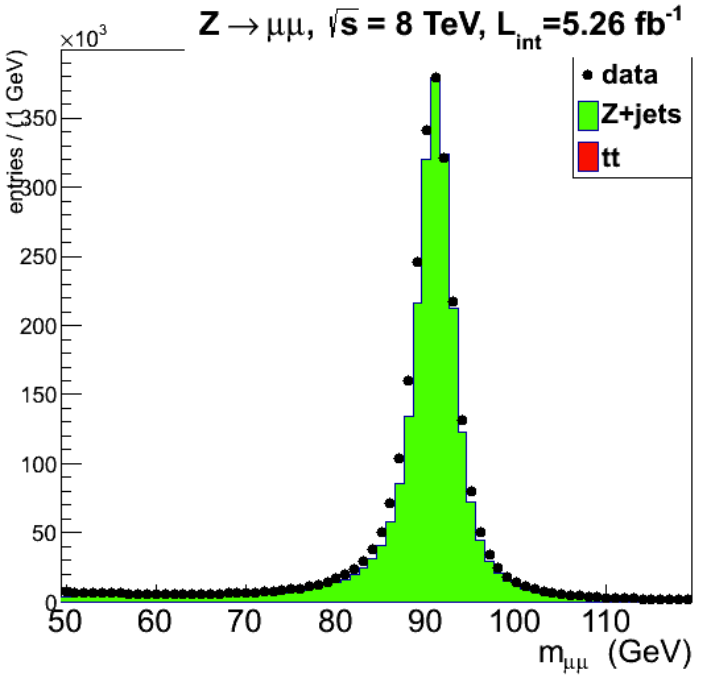


Estimate of sytematics using $Z \rightarrow l^+l^-$

- The not-perfect matching between data and MC highlights the presence of **systematics**:
  - Lepton ID, isolation and SIP
  - Trigger
  - Integrated luminosity
  - Detector calibration
- Better agreement in 7 TeV data

Mismatch data-MC	7 TeV	8 TeV
Z→μμ	1.9%	9.2%
Z→ee	0.8%	0.7%



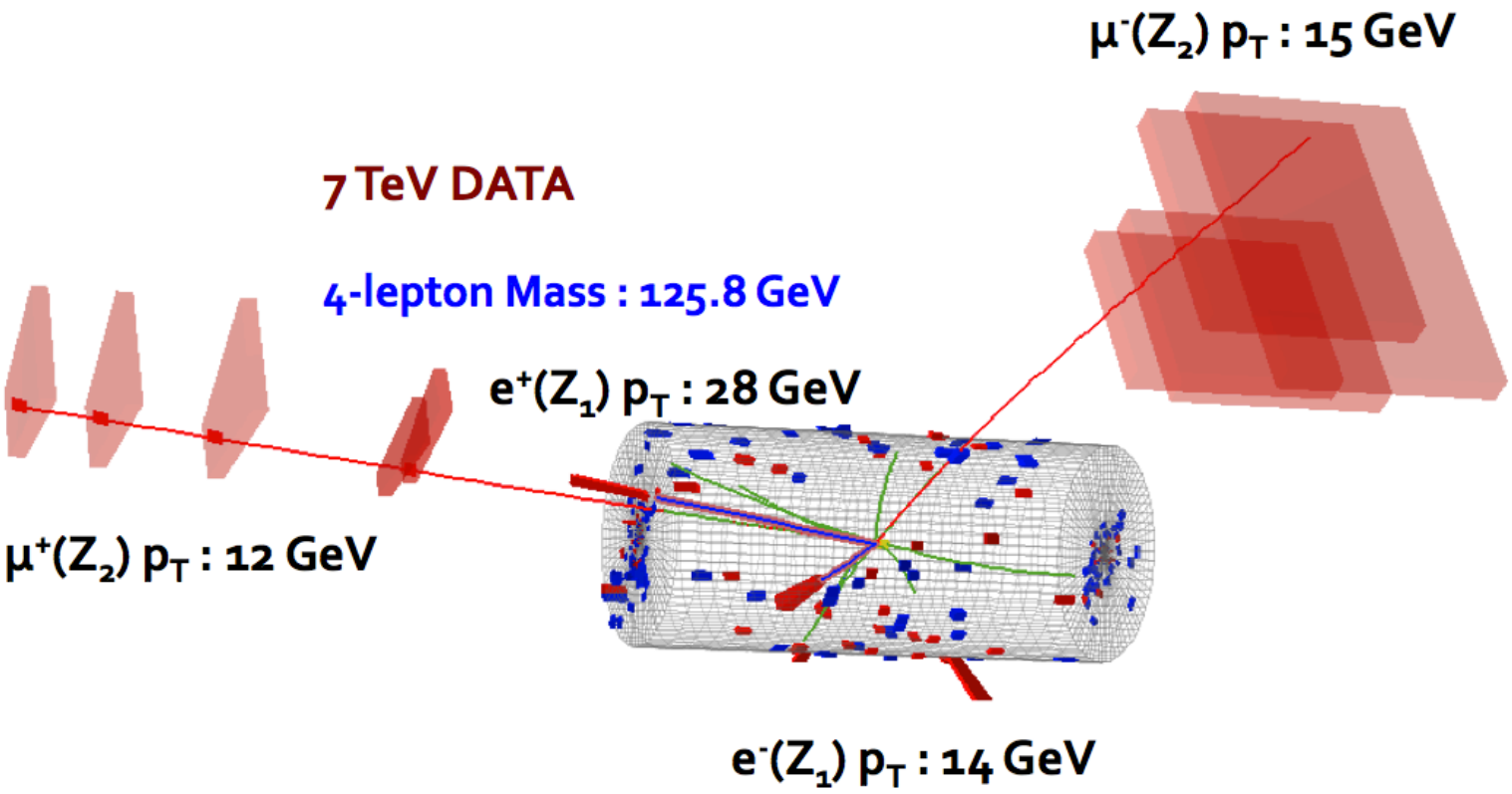


The H→ZZ→4l analysis: **Event selection**

Trigger
double lepton trigger
cross triggers

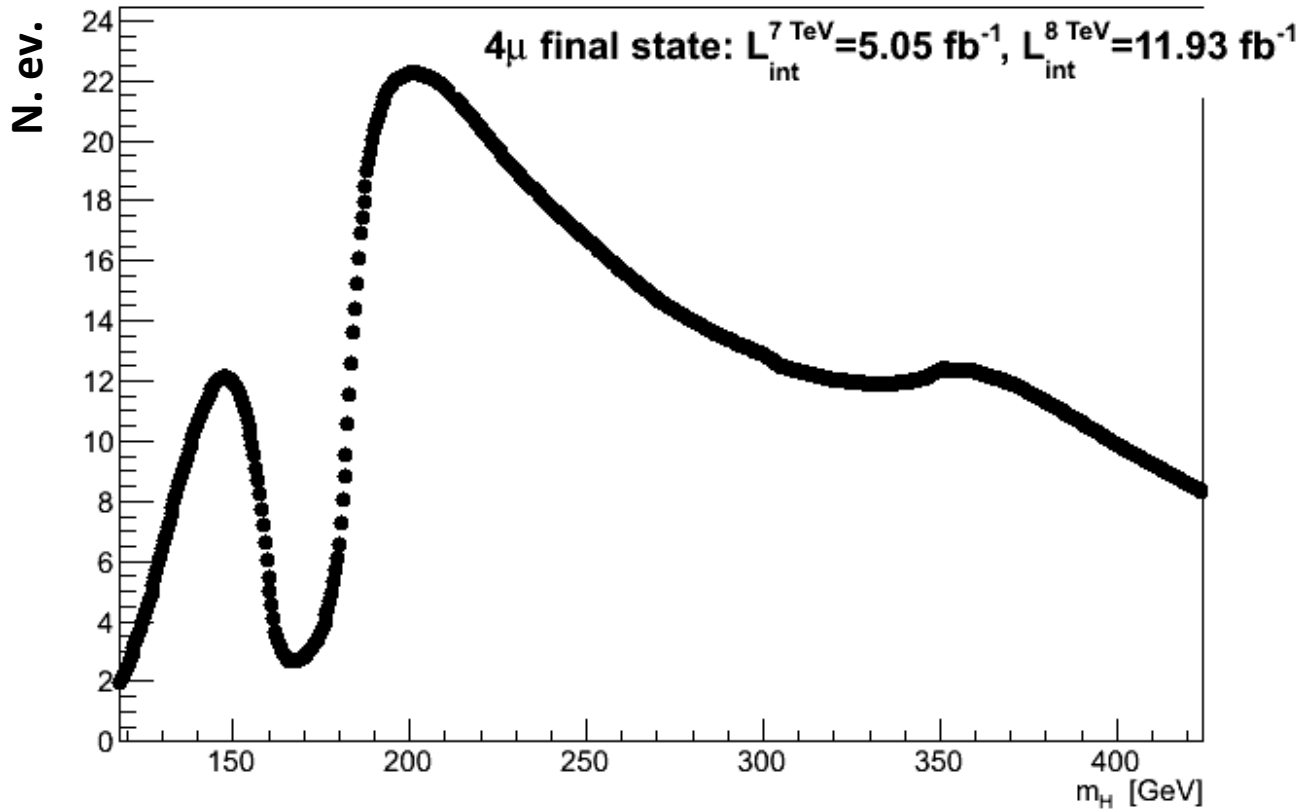
Lepton selection
ID criteria
Isolation and SIP

Build the Z candidates
2 lepton pairs
with FSR recovery



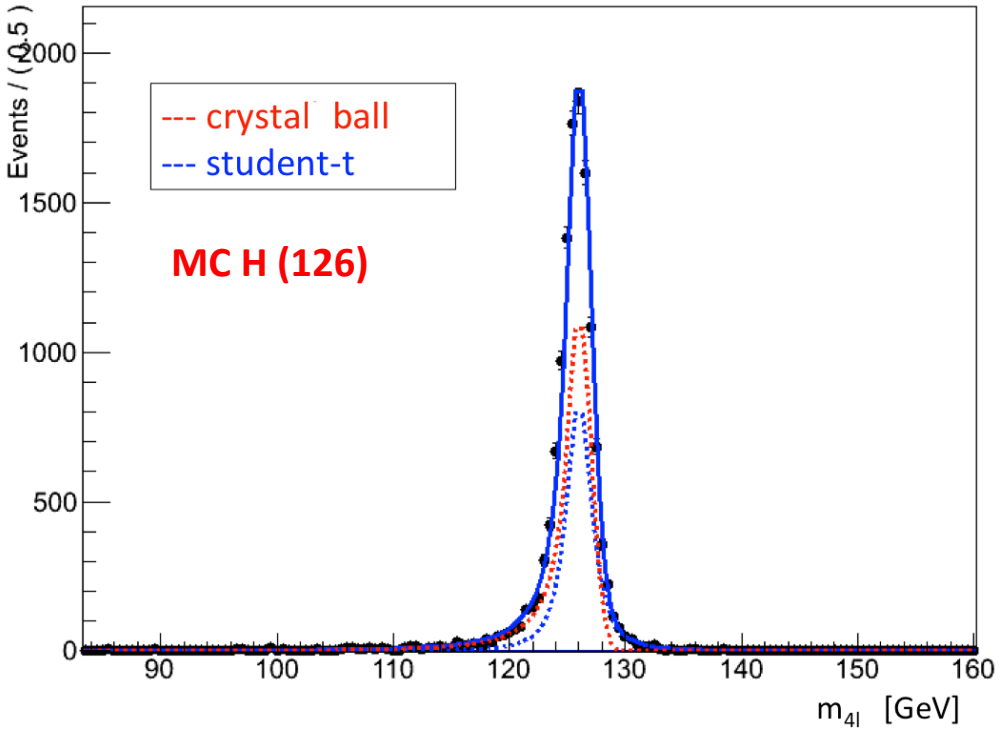
Signal modeling

- The signal yields are estimated directly from theory + MC
  - for a low mass Higgs O(10) events are expected to pass the final selection (summing all the final channels 4μ 4e 2e2μ)
- Systematic uncertainties on cross section
- Systematics on efficiency → from $Z \rightarrow l^+l^-$ systematics



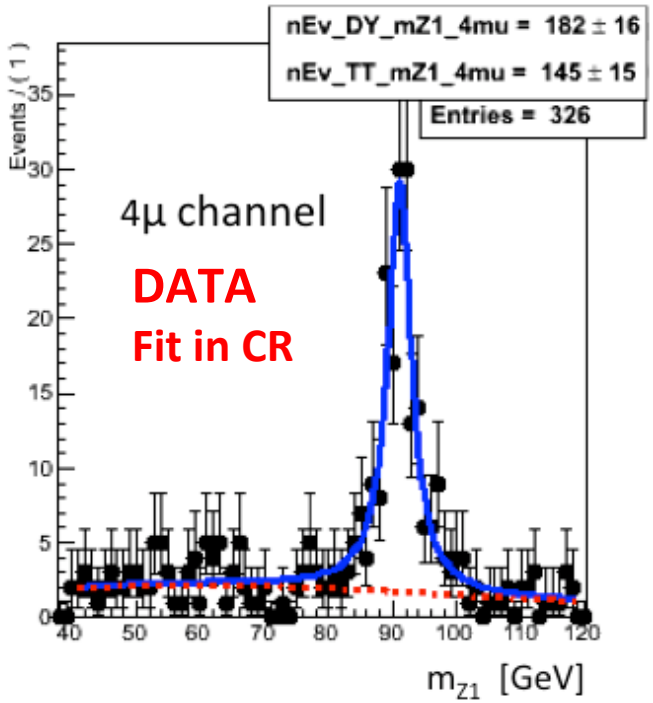
Signal modeling

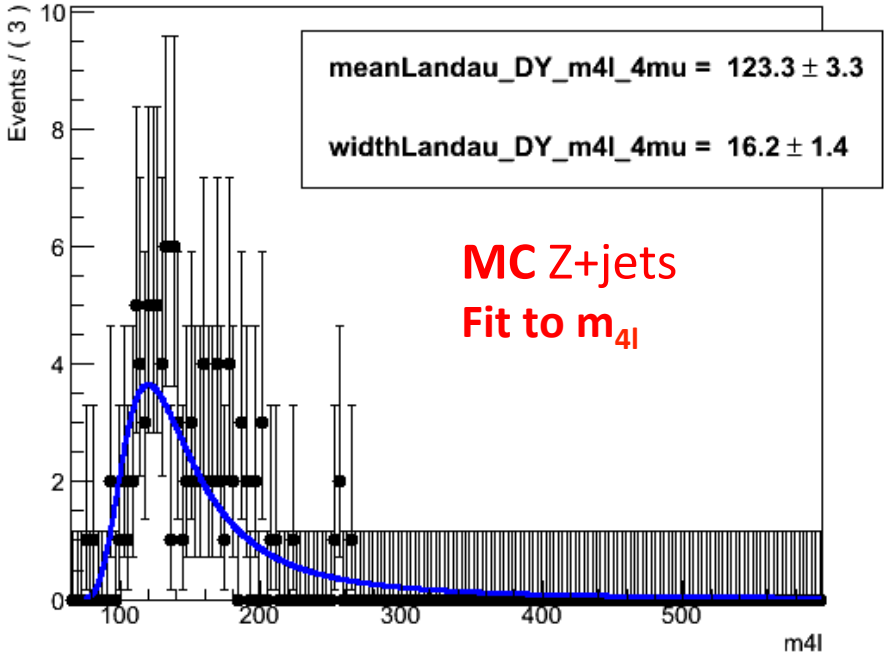
- The peaking mass distribution can be modeled in many ways
  - Chosen an effective parameterization, valid at low masses:
    Student-t + Crystal-Ball for the left tail
- The same pdf is used for Higgs masses from m_H~115 GeV to m_H~180 GeV, the pdf shape is propagated making its parameters mass-dependent



Reducible background (Z+jets, tt)

- **Data-driven method**:
  - **Control Region (CR)** defined inverting the cuts on Isolation/SIP
  - Data collected in the CR are **fitted** to extract the Z+jets and tt yields
  - Extrapolation to Signal Region (SR) using MC
  - Mass shapes from MC
- Data yield in the CR are well understood

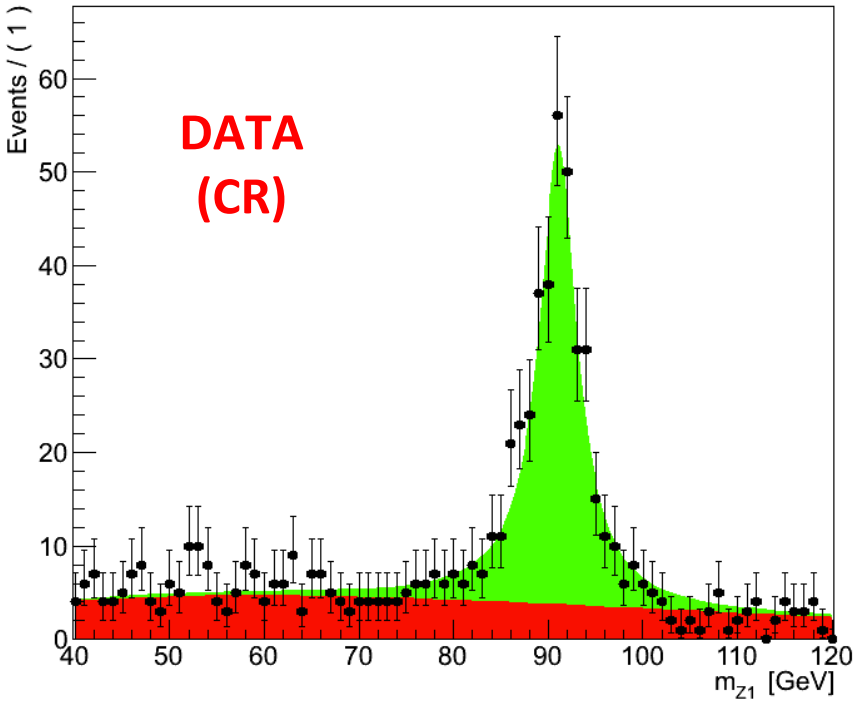


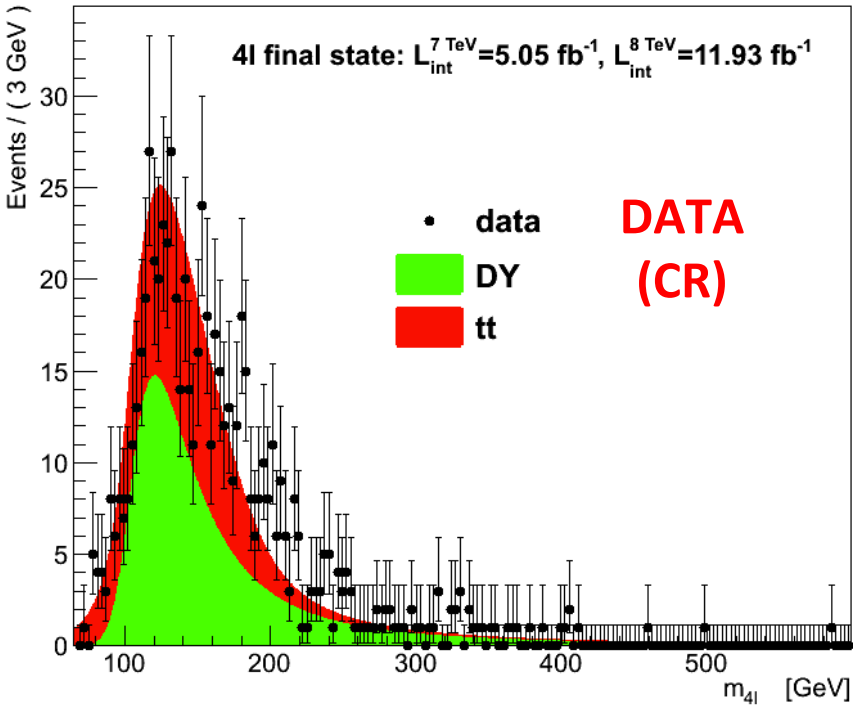


Reducible background

- Data-driven method:
  - **large uncertainty** on yields, dominated by low-MC statistics
  - still **acceptable results**, minor degradation of final sensitivity

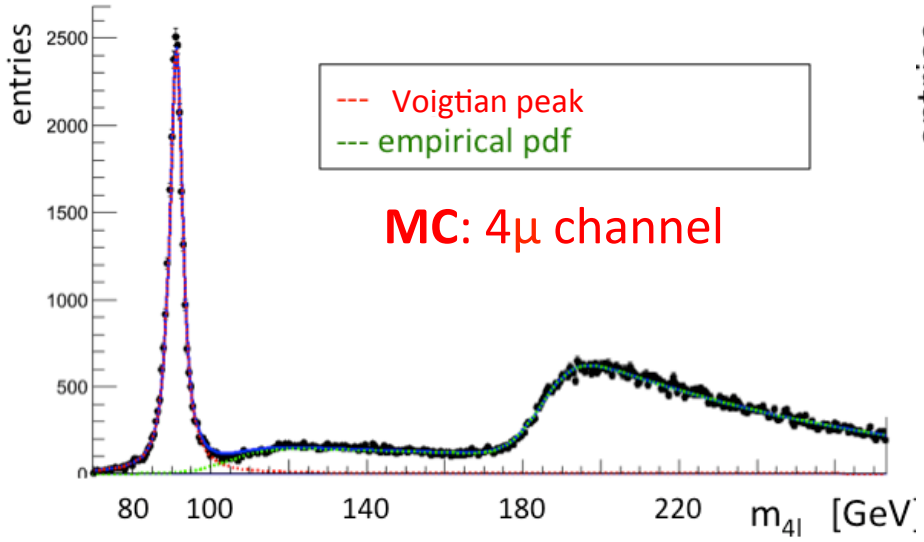
Process	N_{SR} (rel.err.)
Z+jets - 4μ	3.3 ± 1.9 (59%)
Z+jets - 4e	2.0 ± 2.1 (105%)
Z+jets - 2e2μ	1.7 ± 1.2 (72%)
tt - 4μ	0.1 ± 0.1 (59%)
tt - 4e	0.2 ± 0.1 (57%)
tt - 2e2μ	0.7 ± 0.2 (29%)

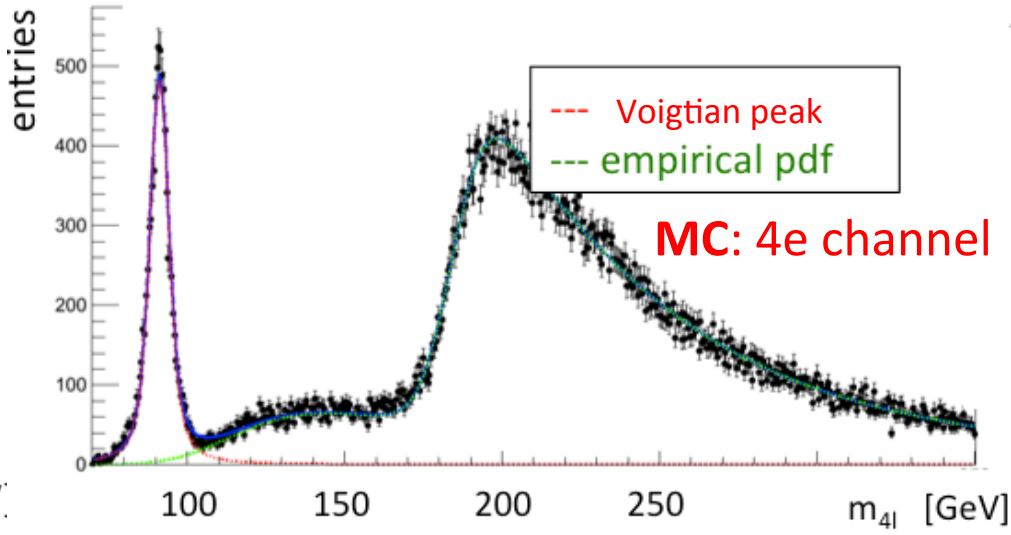




Irreducible background

- qq→ ZZ/Zγ gg→ZZ ZZ→2l2τ
- The pure predictions from theory-MC are reliable
- Mass shape
  - empirical fit to the ZZ continuum
  - Voigtian peak for the Z→4l
  - Landau convoluted with a Gaussian for ZZ→2l2τ

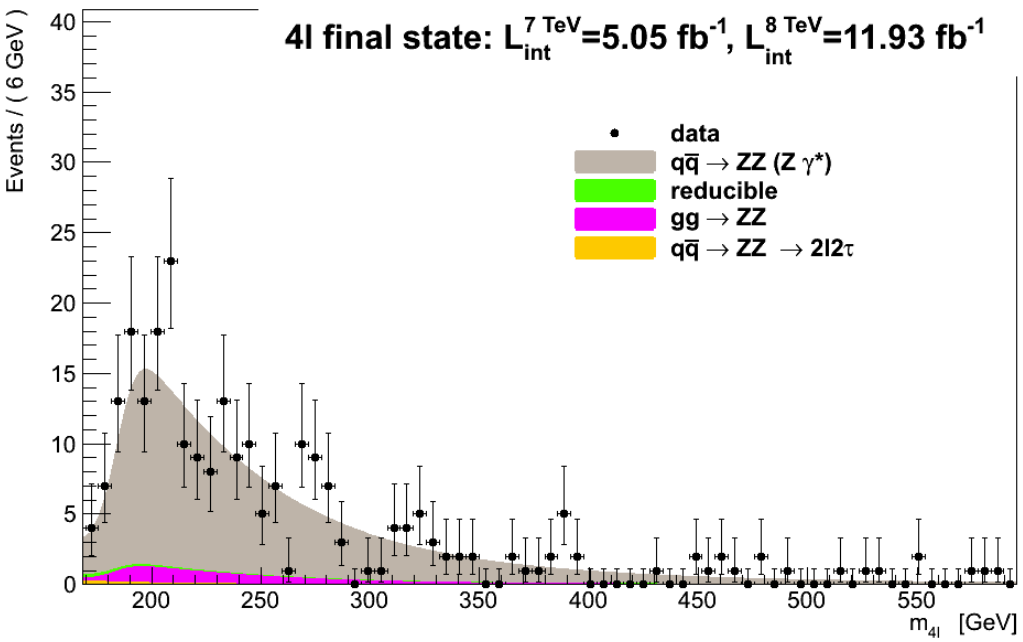




Irreducible background

- A **Data-driven technique** is used to control the yields from data
  - **Renormalization Band (RB)** for m_{4l} > 180 GeV (signal free region)
  - Global yield in SR renormalized to the number of events found in the RB

RB	4μ	4e	2e2μ	TOT
N^{MC}_{RB}	70.36	48.91	109.18	228.45
N^{DATA}_{RB}	71	46	129	246



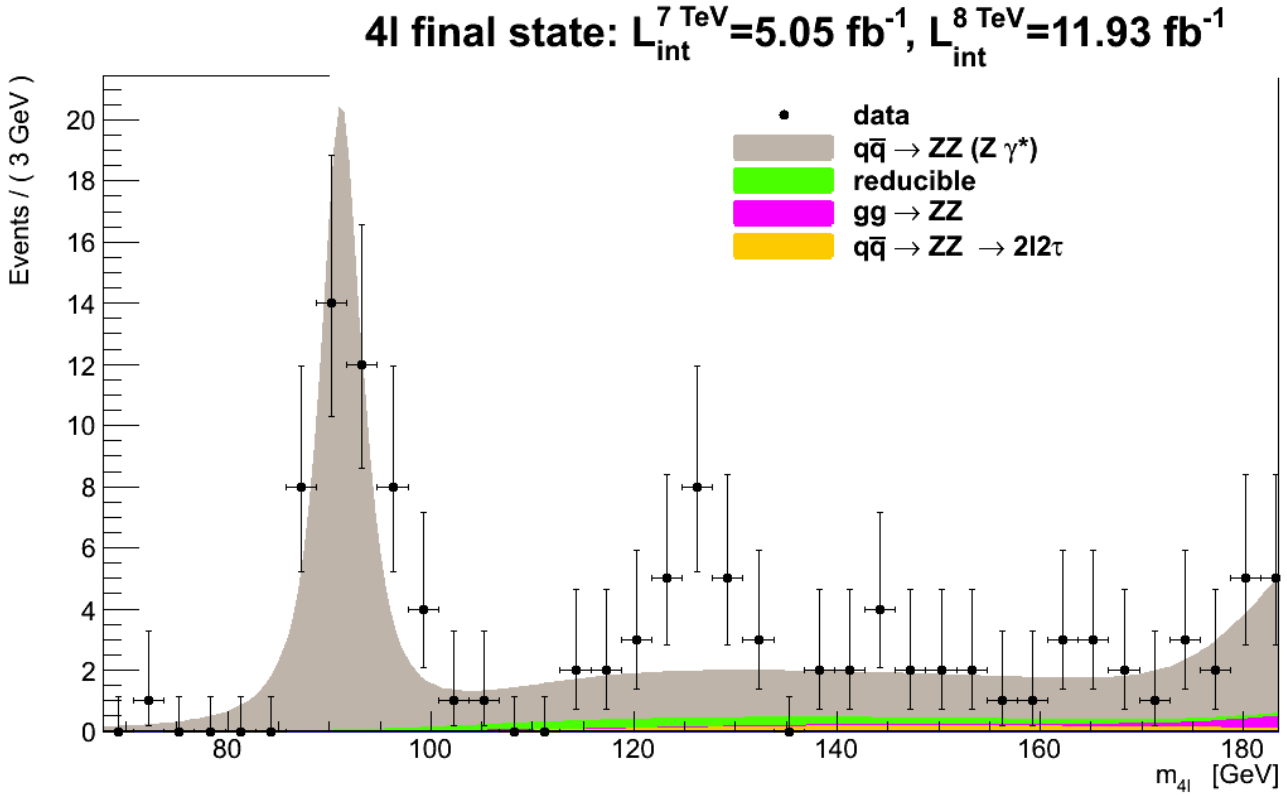
Global yield SR	ZZ/Zγ	gg→ZZ	ZZ →2l2τ
4μ	98.4 ± 11.7	5.3 ± 0.6	1.5 ± 0.2
4e	52.1 ± 7.7	3.6 ± 0.5	1.0 ± 0.1
2e2μ	155.6 ± 13.7	10.7 ± 0.9	2.1 ± 0.2
TOT	306.2 ± 19.6	19.6 ± 1.2	4.6 ± 0.3

With DD technique: Statistics uncertainties (of the same order of systematic ones, that are now removed)

Results

- **Data excess peaking around 126 GeV** over the almost flat background
- A statistical analysis is necessary to evaluate its significance

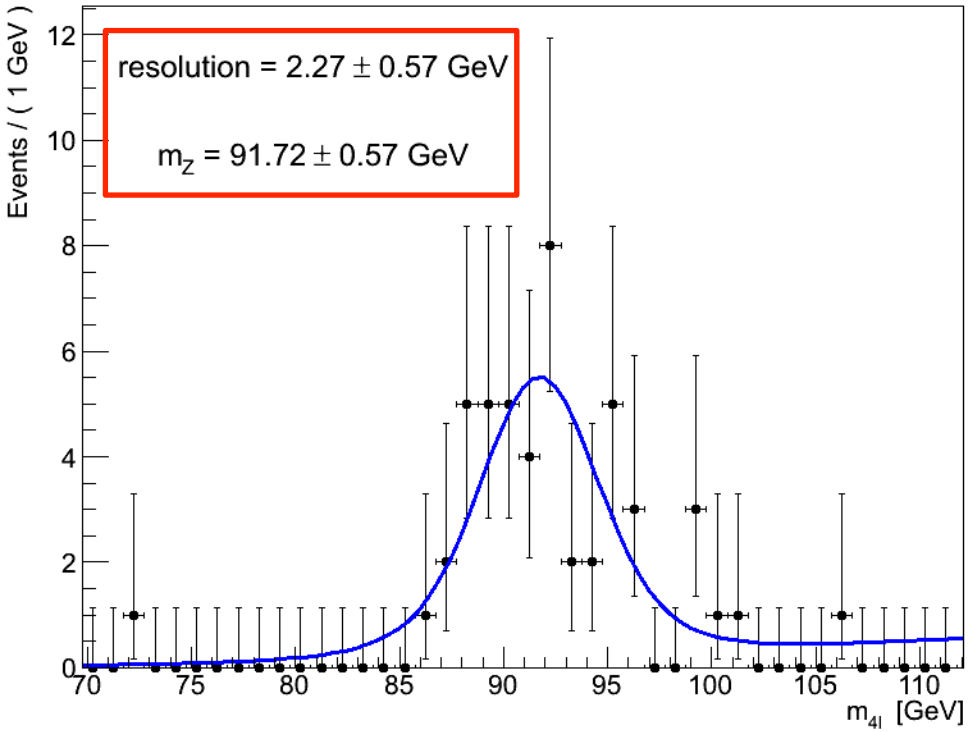
[110,180]	4μ	4e	2e2μ	TOT
ZZ	11.2	5.5	19.3	36
ggZZ	0.4	0.2	0.9	1.6
2l2τ	0.9	0.5	1.8	3.3
Red.	2	1.4	1.6	5.5
Total bkg	14.5	7.7	23.7	45.8
Obs	**20**	**14**	**26**	**60**



The Z→4l peak

- The Z→4l peak is a benchmark to test the capability to correctly reconstruct a 4l-peak
  - we fit correctly its position, and we have an estimate of 4l-mass resolution

[70,100] GeV	4μ	4e	2e2μ	TOT
MC	20	5.8	15.8	41.6
DATA	30	8	9	47



Statistical analysis – the CLs method

- Statistical analysis performed with the **profile likelihood method**:
  - **Observable: m_{4l}** → **distributions specified under S+B and B hypotheses**
  - **Parameter of interest: $\mu = XS/XS_{SM}$**
  - Modeling of systematic uncertainties θ

→ *Hat (^) if fitted to data*

- **Significance of an excess**:

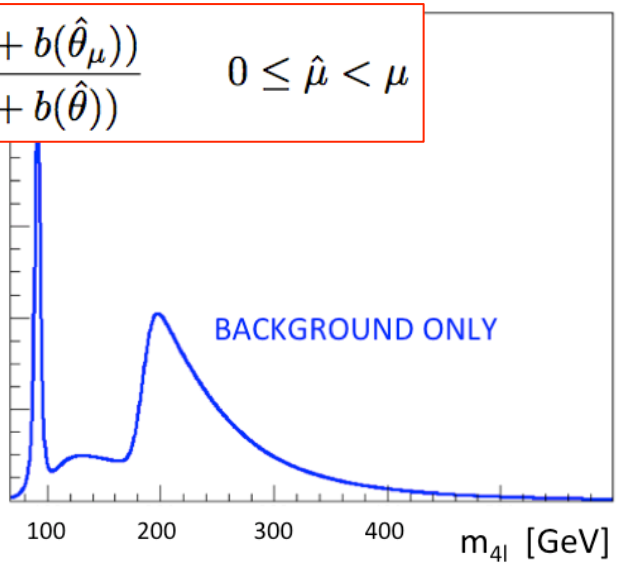
$$q_0 = -2\log\frac{L(data|b(\hat{\theta}_0))}{L(data|\hat{\mu}\cdot s(\hat{\theta})+b(\hat{\theta}))} \qquad \hat{\mu}\geq 0$$

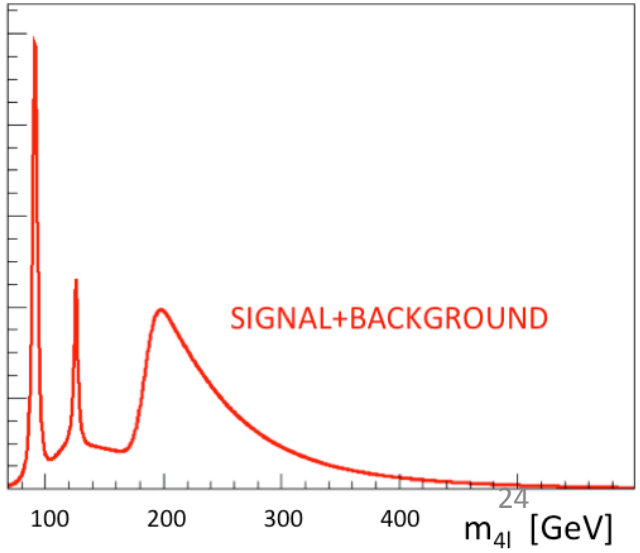
$$p_0 = P(q_0 \geq q_0^{obs}|b)$$

- **95% CL upper limits → exclusions**:

$$q_\mu = -2\log\frac{L(data|\mu\cdot s(\hat{\theta}_\mu)+b(\hat{\theta}_\mu))}{L(data|\hat{\mu}\cdot s(\hat{\theta})+b(\hat{\theta}))} \qquad 0\leq\hat{\mu}<\mu$$

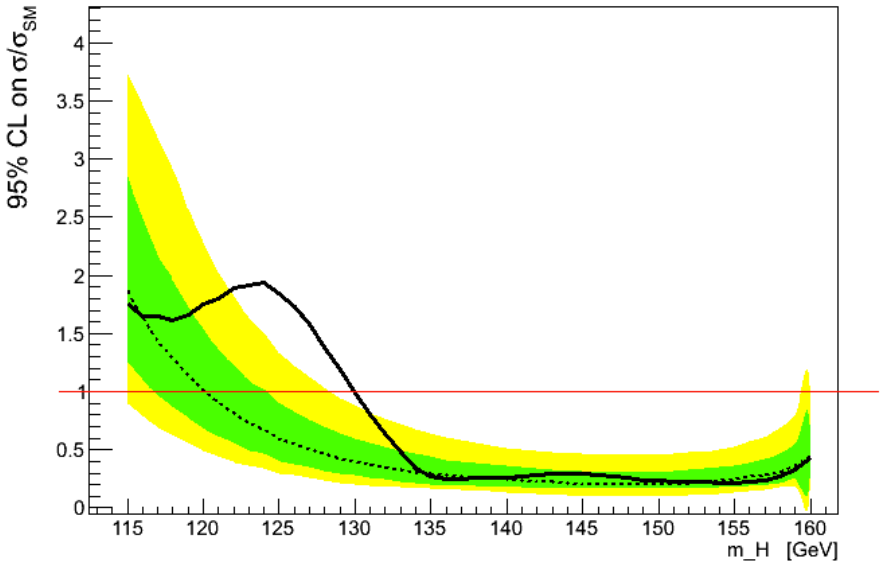
$$CL_s = \frac{CL_{s+b}}{CL_b}$$

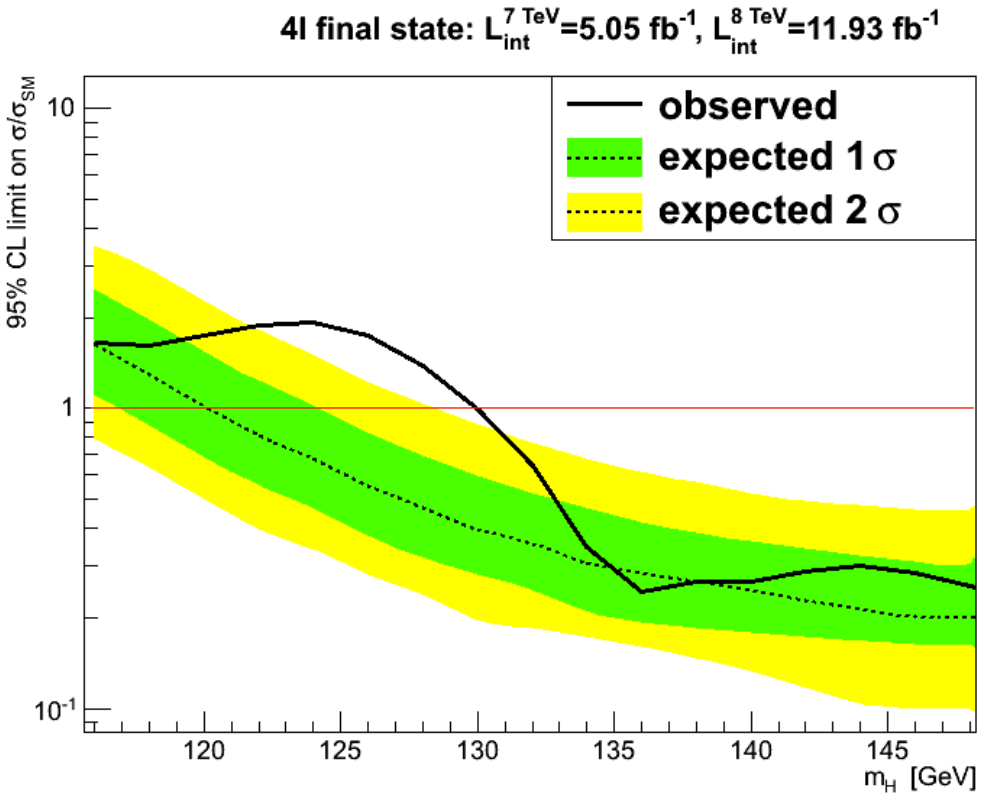




Statistical analysis – Upper limits

- **Excluded mass range** at 95% CL: **> 130 GeV**
- Excess wrt background at low mass prevents to exclude the signal here

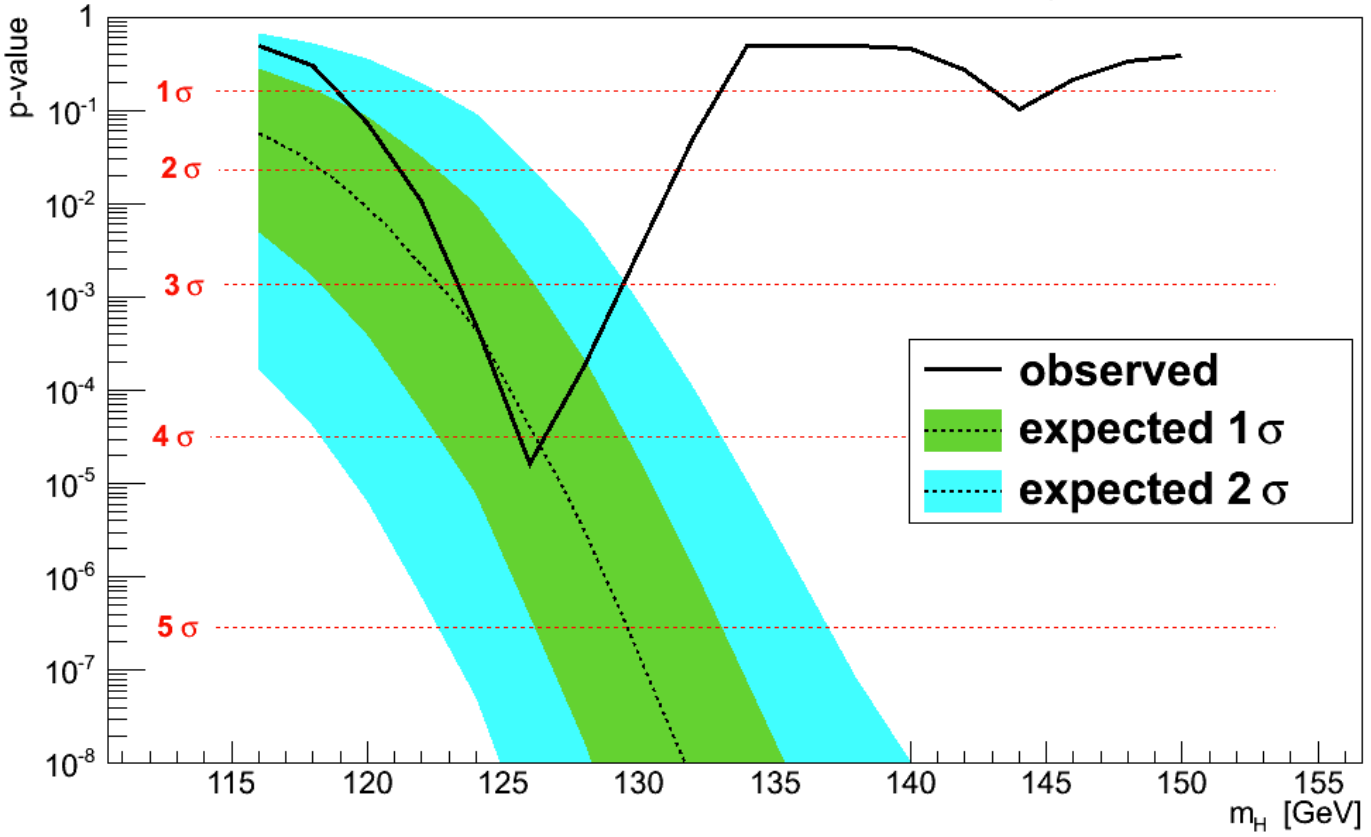




Statistical analysis – Results

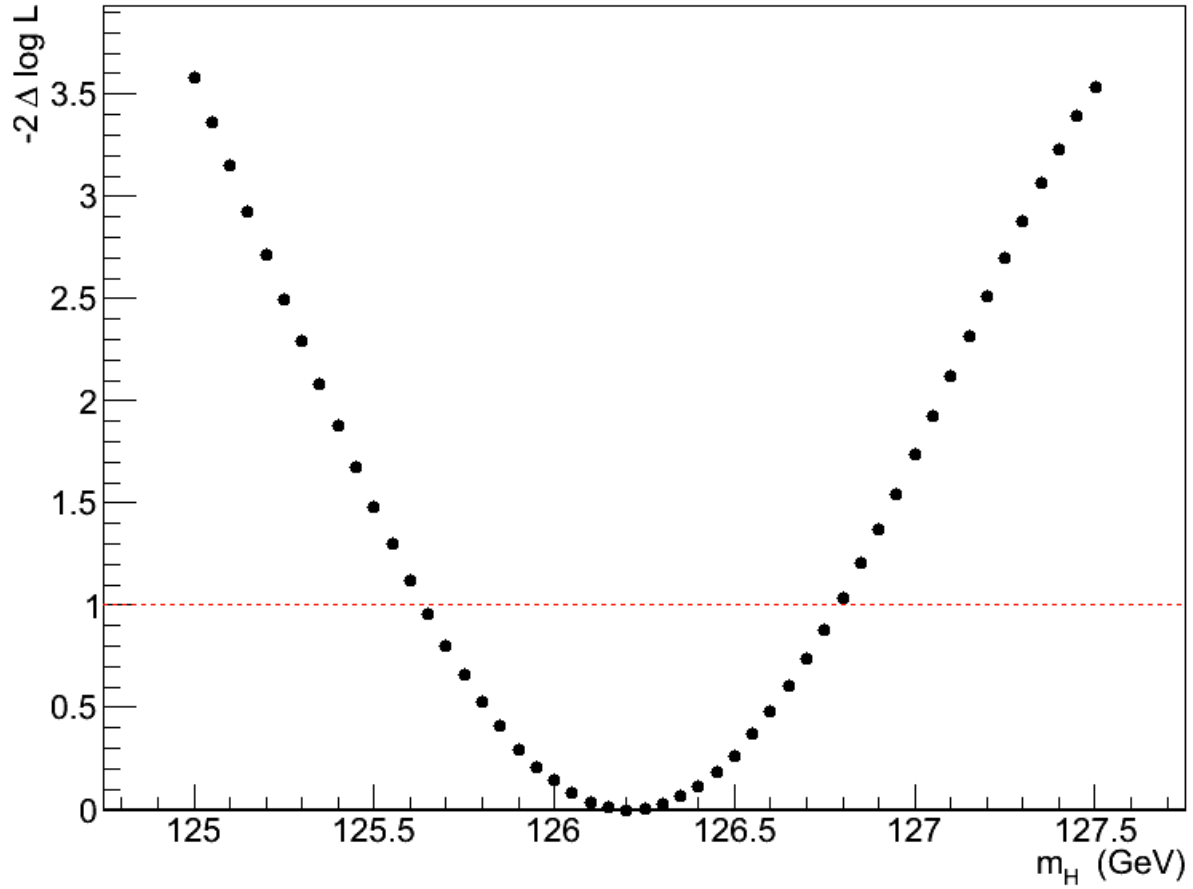
- Largest excess wrt backround at m_H = 126 GeV: **local significance 4.1σ**

Evidence!



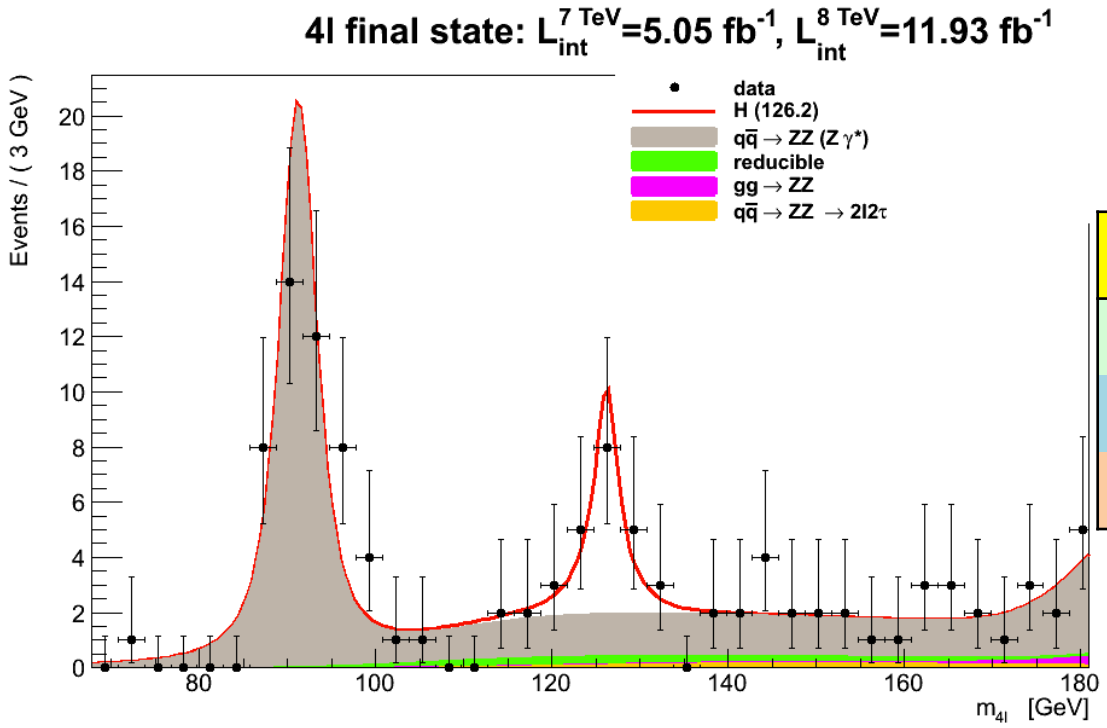
Measurement of properties

- Fit to **cross section**: $\mu = \sigma/\sigma_{SM} = 1.03 \pm 0.36$
- Graphical fit to the **mass**: $m_H = 126.2 \pm 0.55$ GeV



Final Results

- The event distribution is in good agreement with SM expectations
- **Peak around 126 GeV** over the almost flat background, compatible with what expected for the SM Higgs



[110,140]	4μ	4e	2e2μ	TOT
Bkg	6.6	2.8	9.6	18.9
Signal	4.5	2.1	6.2	12.8
Obs	**11**	**9**	**10**	**30**

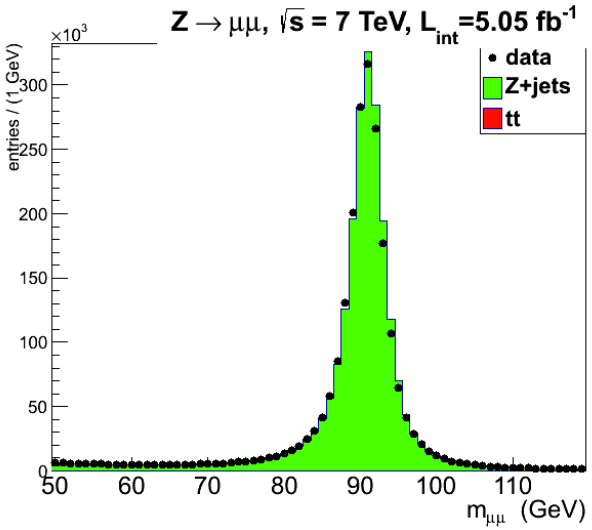
Conclusions

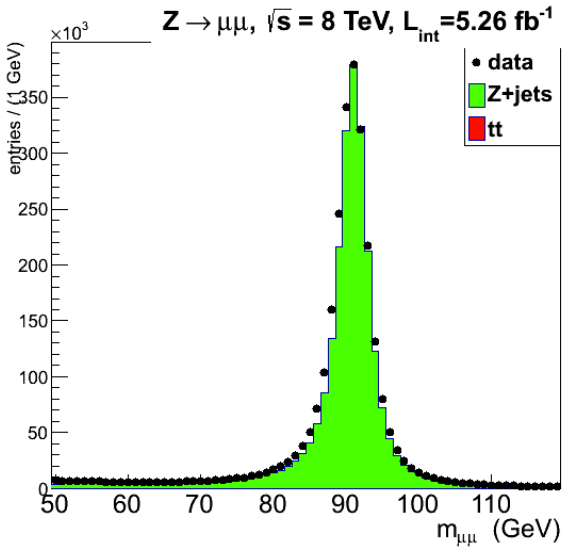
- The results for the SM Higgs search at CMS in the H→ZZ→4l search channel have been presented. This search makes use of the excellent detector performances in reconstructing physical objects
- Data-driven techniques to control the background help to avoid biases
- The search sees an **excess** in data **compatible with the Standard Model Higgs boson → local significance 4.1 standard deviations**
- The fitted **cross section** is **1.03 ± 0.36 the standard model prediction**
- The fitted **mass** is **126.2 ± 0.55 GeV**
- The analysis can be considered a cross-check of the CMS official one
- More precise results require more data:

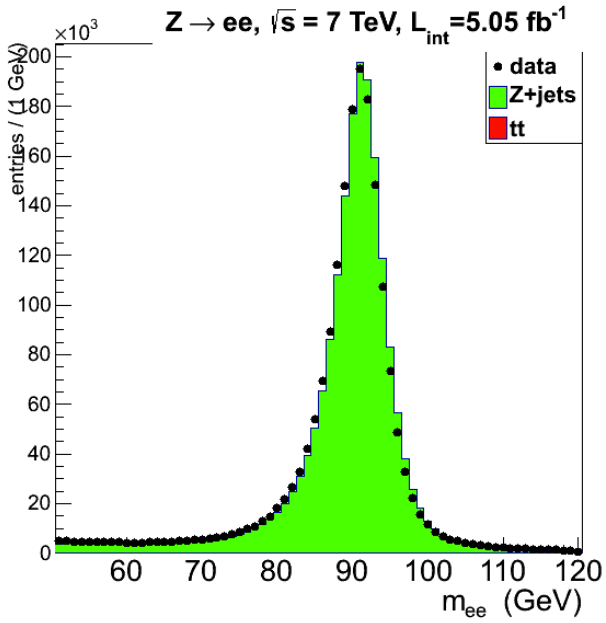
  22 fb^{-1} collected by the end of 2012 at 8 TeV

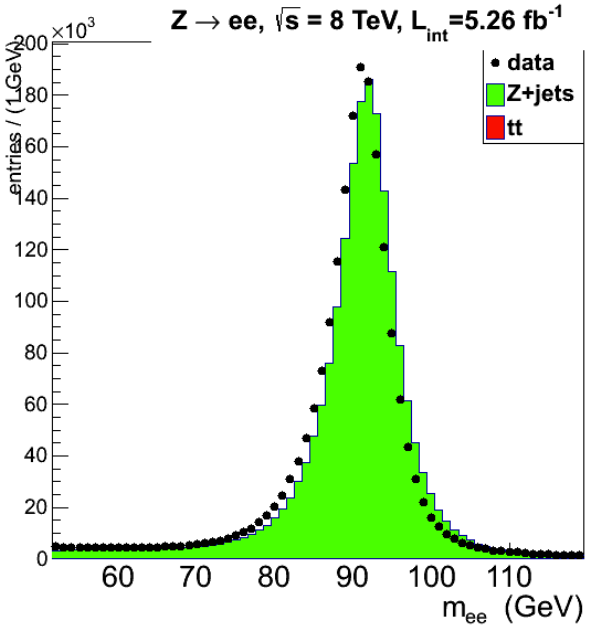
Backup

Z→ll

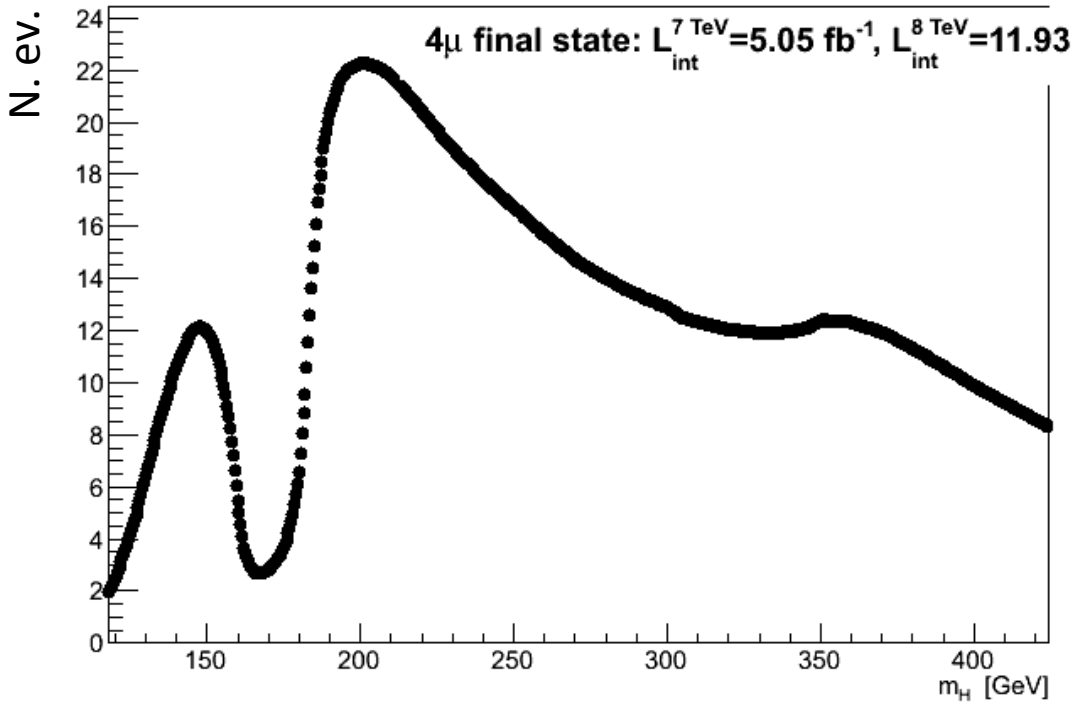

Z → μμ, √s = 7 TeV, L_int=5.05 fb^-1
×10^3
entries / (1 GeV)
data
Z+jets
tt
m_μμ (GeV)

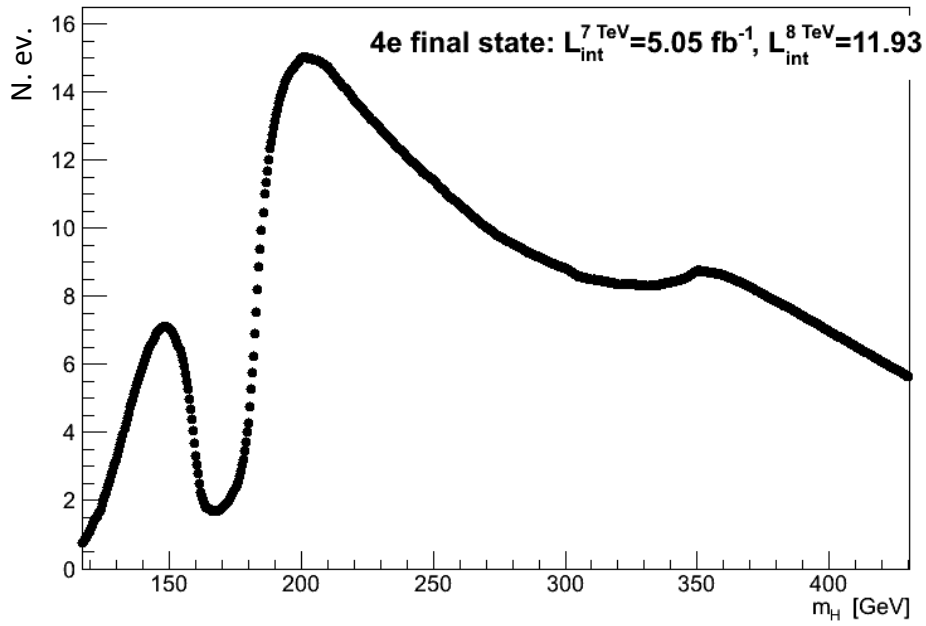

Z → μμ, √s = 8 TeV, L_int=5.26 fb^-1
×10^3
entries / (1 GeV)
data
Z+jets
tt
m_μμ (GeV)

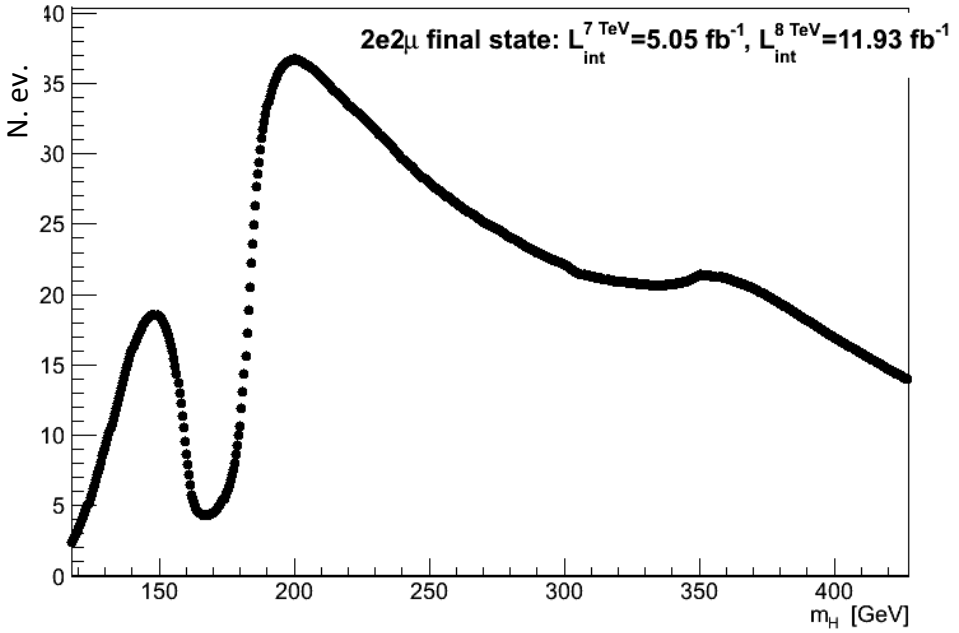

Z → ee, √s = 7 TeV, L_int=5.05 fb^-1
×10^3
entries / (1 GeV)
data
Z+jets
tt
m_ee (GeV)


Z → ee, √s = 8 TeV, L_int=5.26 fb^-1
×10^3
entries / (1 GeV)
data
Z+jets
tt
m_ee (GeV)

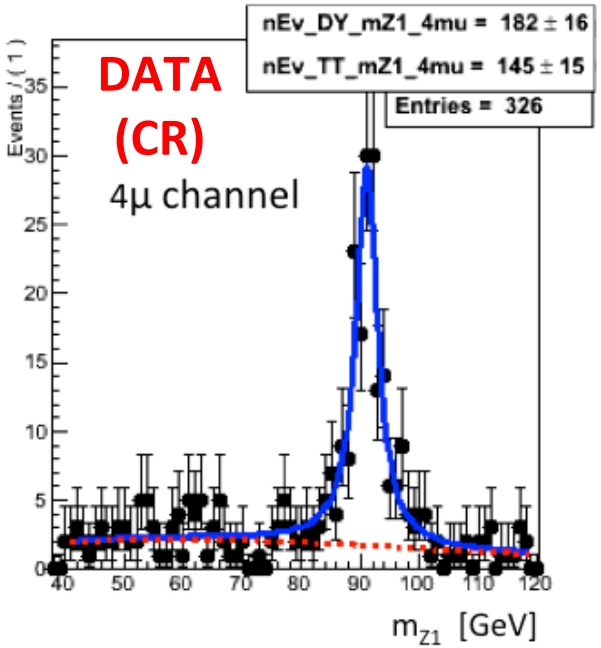
Signal Yields

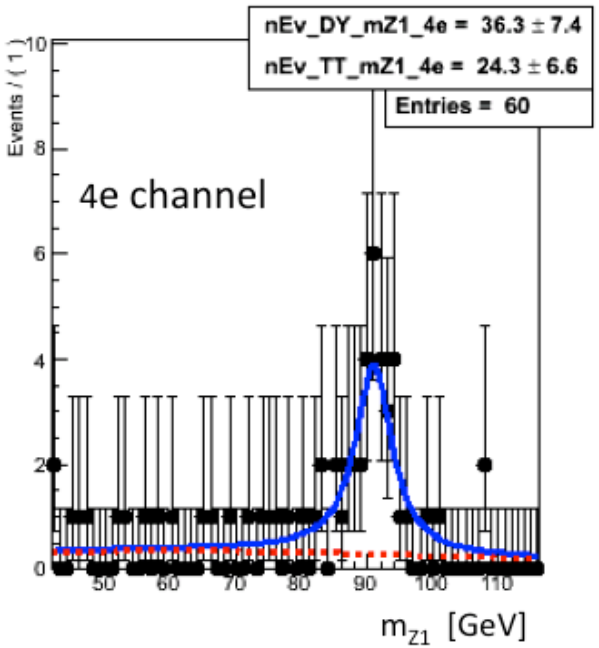


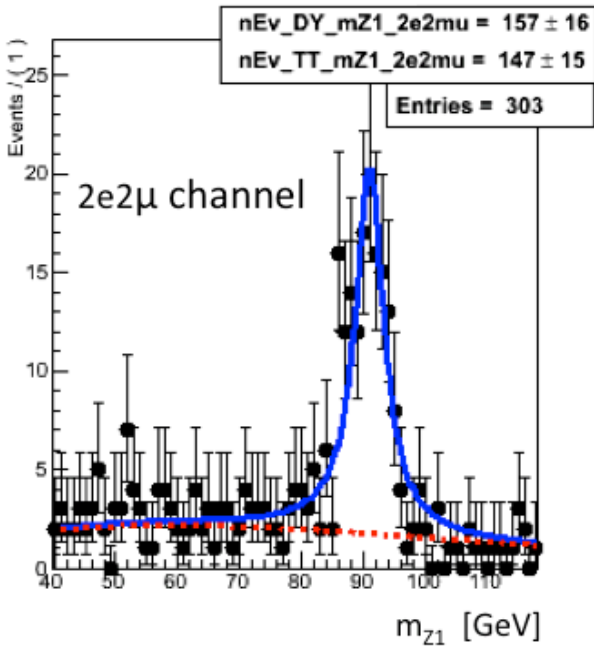


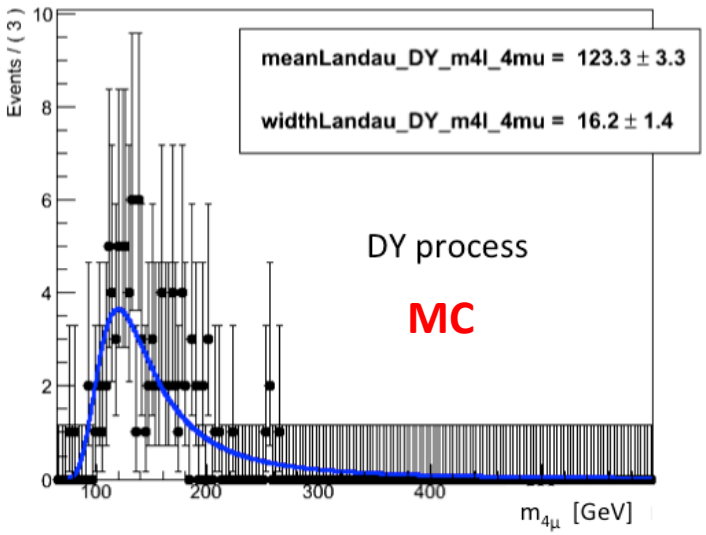


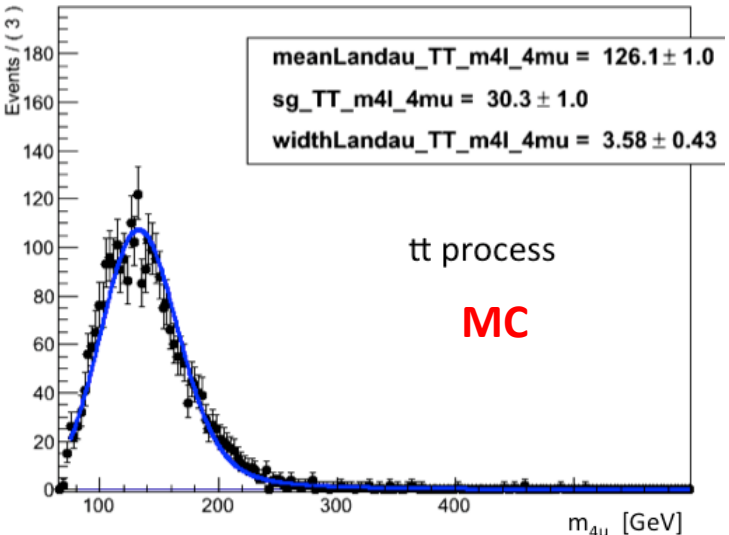
Reducible bkg- Fit in CR



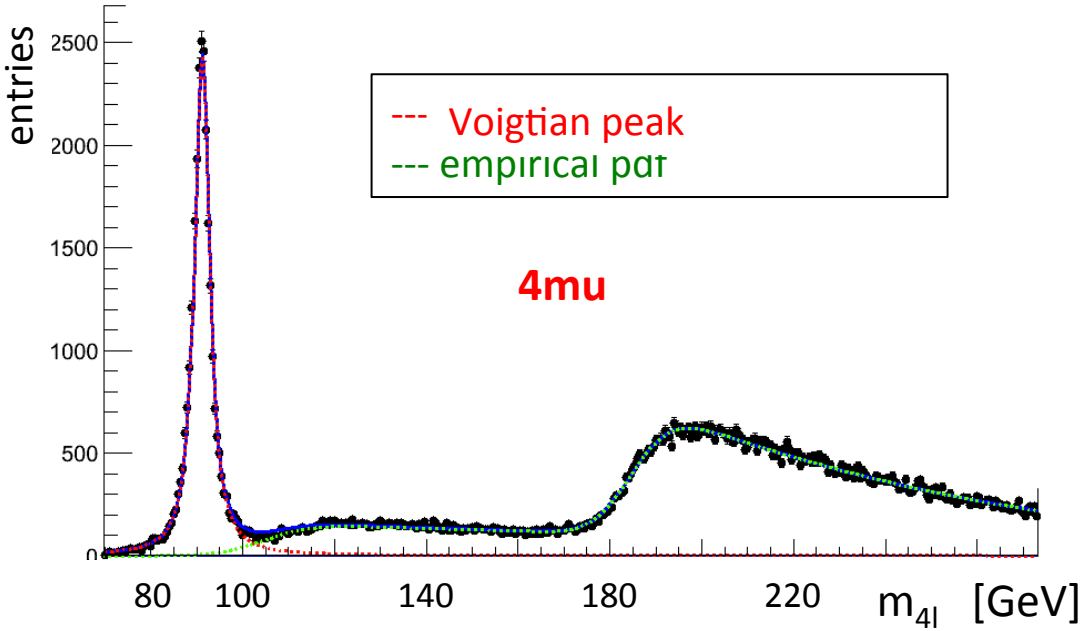


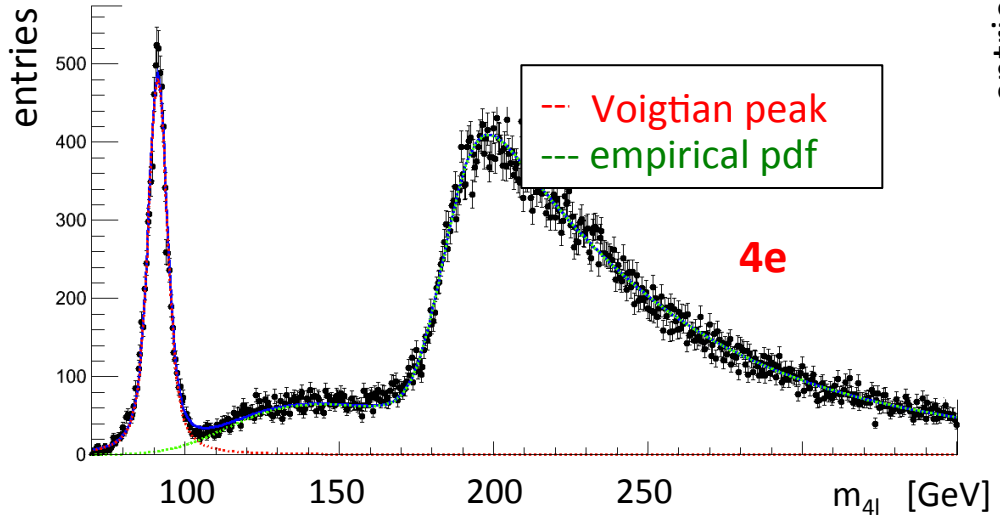


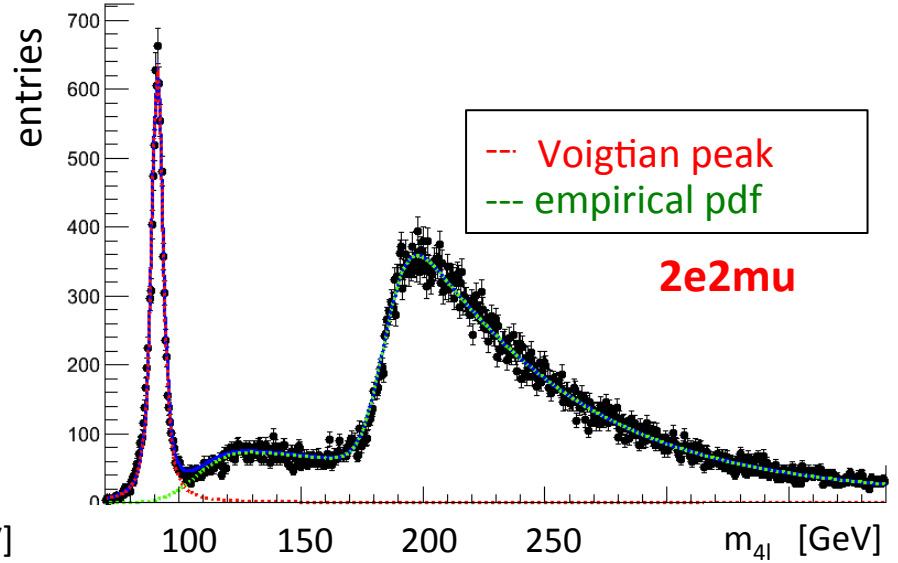




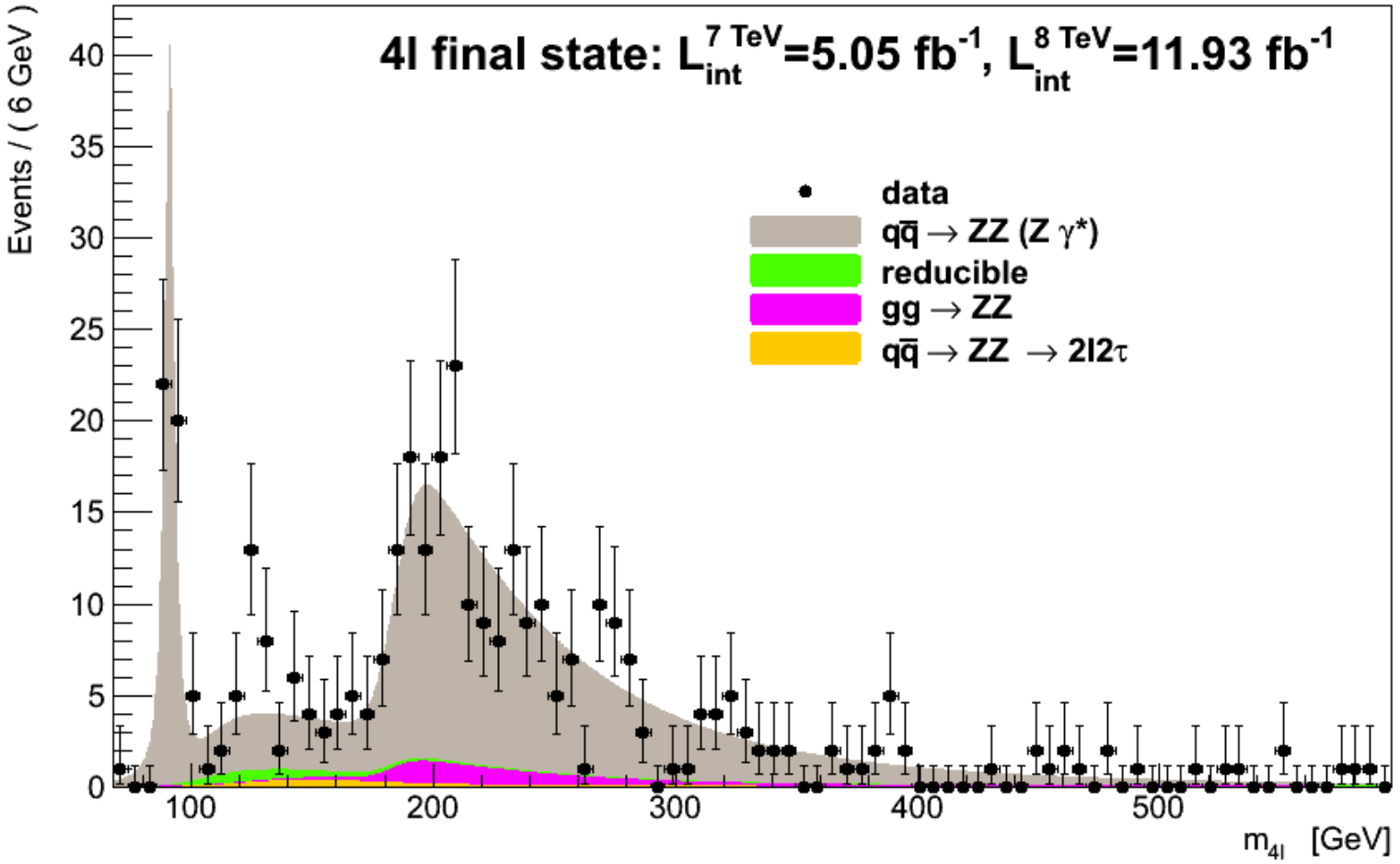
Irreducible bkg



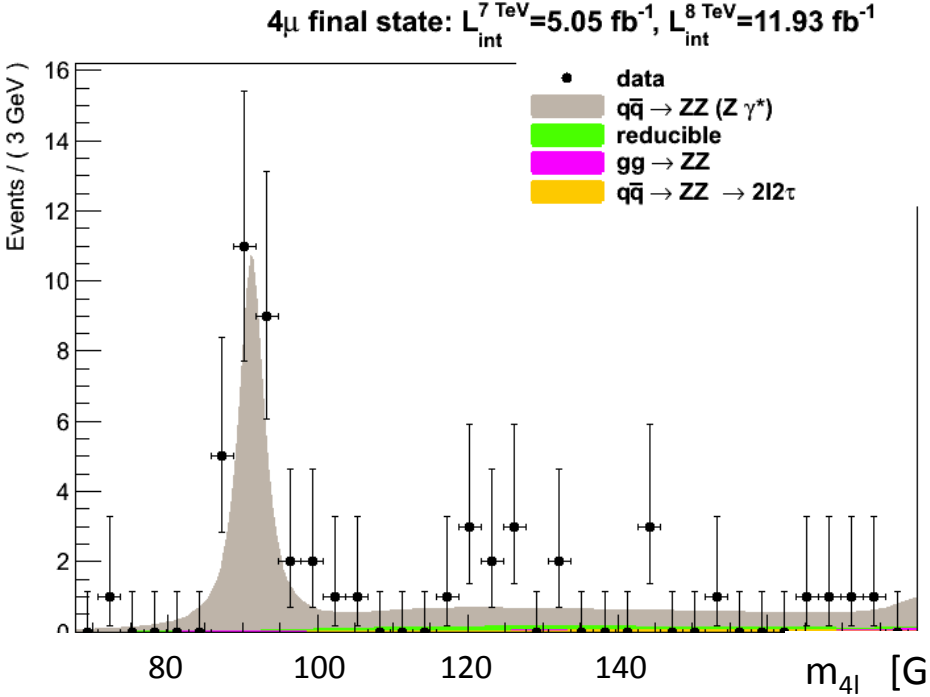


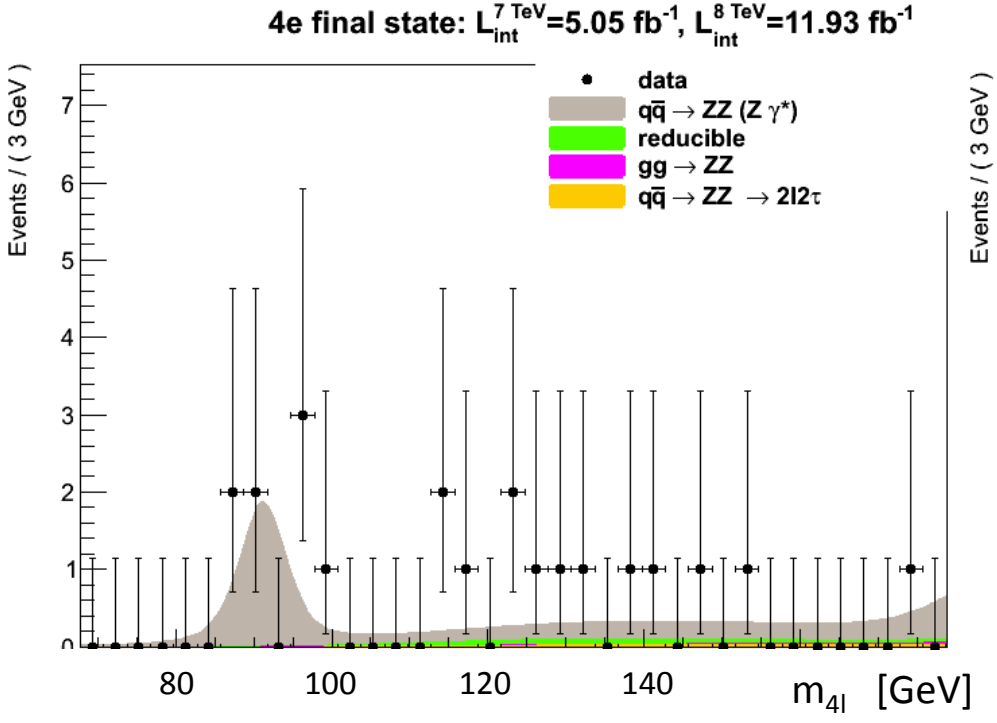


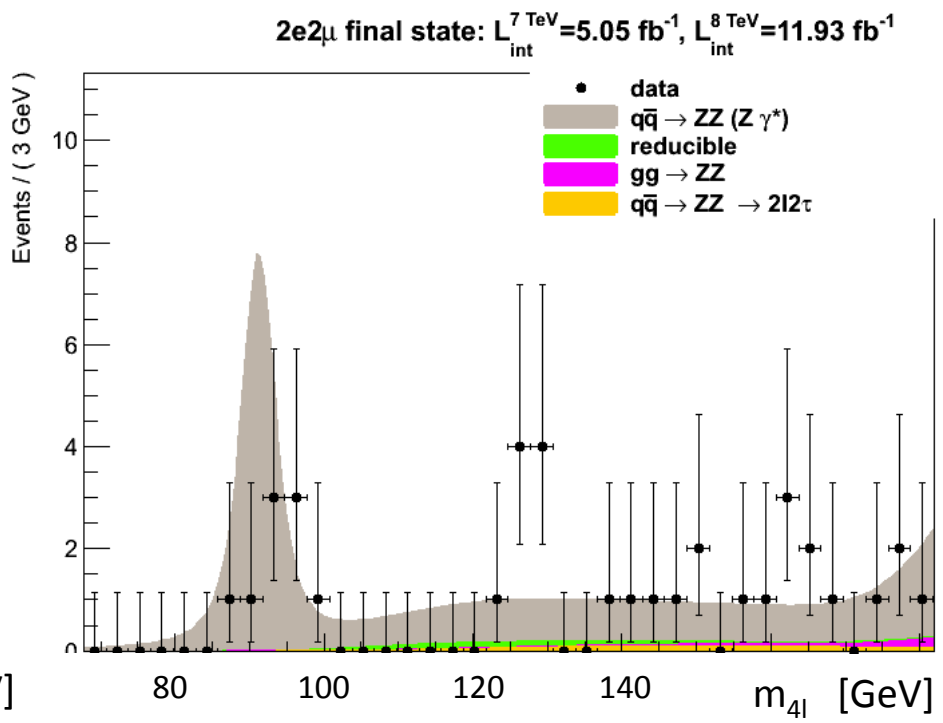
Full mass spectrum



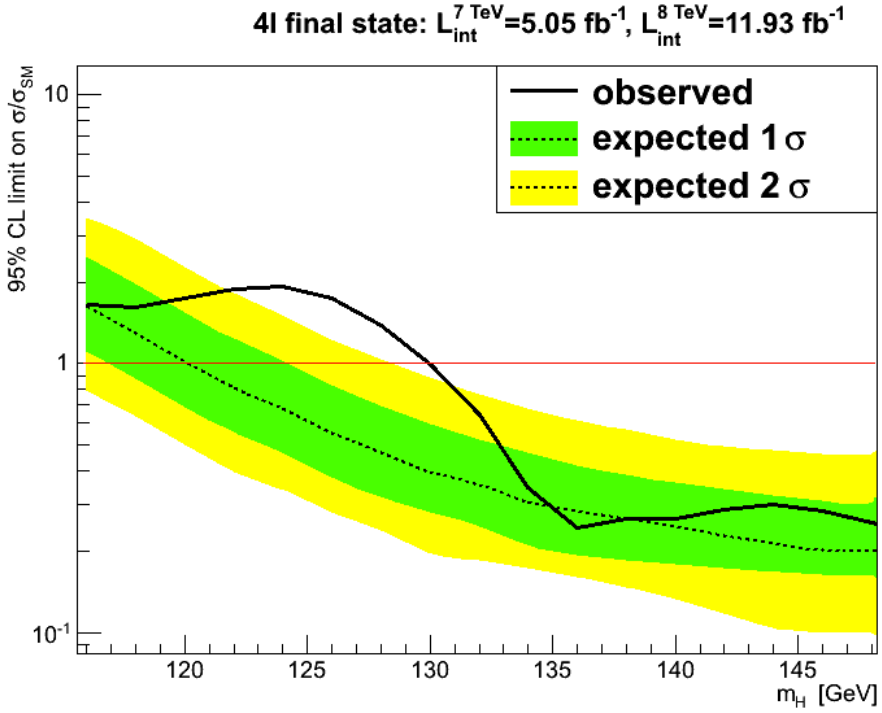
channels

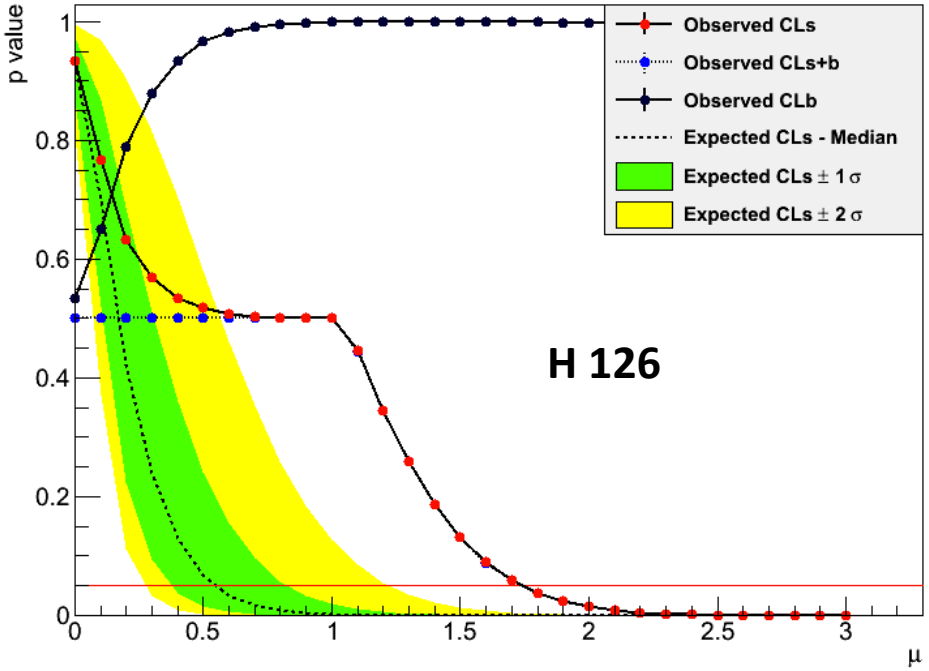

4μ final state: $L_{int}^{7\ TeV}$=5.05 fb^{-1}, $L_{int}^{8\ TeV}$=11.93 fb^{-1}
Events / (3 GeV)
data
qq̄ → ZZ (Z γ*)
reducible
gg → ZZ
qq̄ → ZZ → 2l2τ
m4l [GeV]


4e final state: $L_{int}^{7\ TeV}$=5.05 fb^{-1}, $L_{int}^{8\ TeV}$=11.93 fb^{-1}
Events / (3 GeV)
data
qq̄ → ZZ (Z γ*)
reducible
gg → ZZ
qq̄ → ZZ → 2l2τ
m4l [GeV]

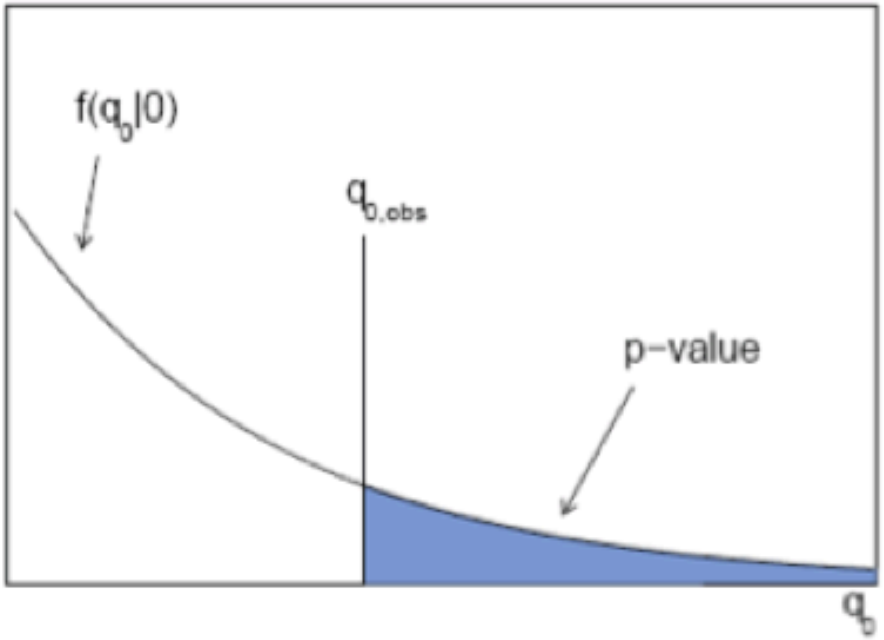

2e2μ final state: $L_{int}^{7\ TeV}$=5.05 fb^{-1}, $L_{int}^{8\ TeV}$=11.93 fb^{-1}
Events / (3 GeV)
data
qq̄ → ZZ (Z γ*)
reducible
gg → ZZ
qq̄ → ZZ → 2l2τ
m4l [GeV]

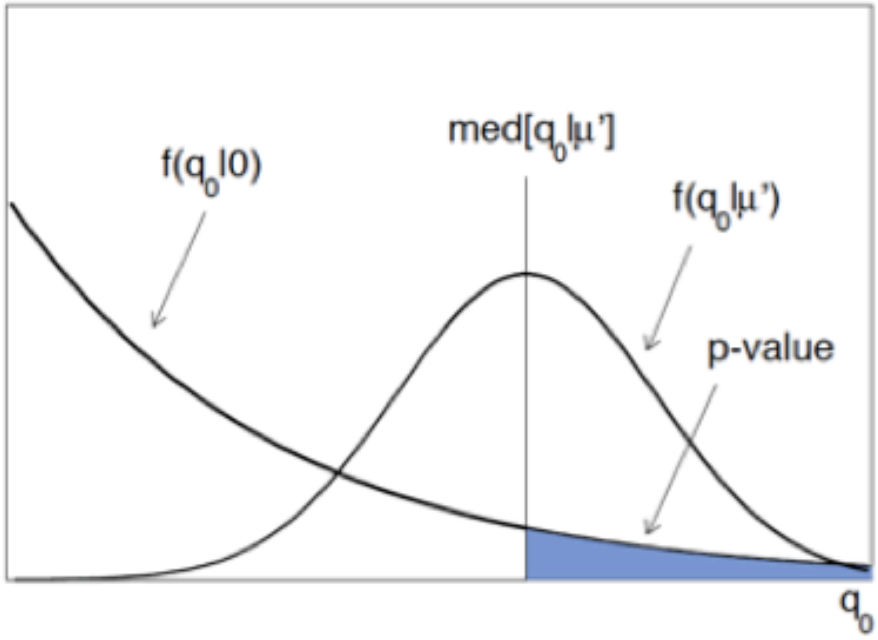
Upper limits



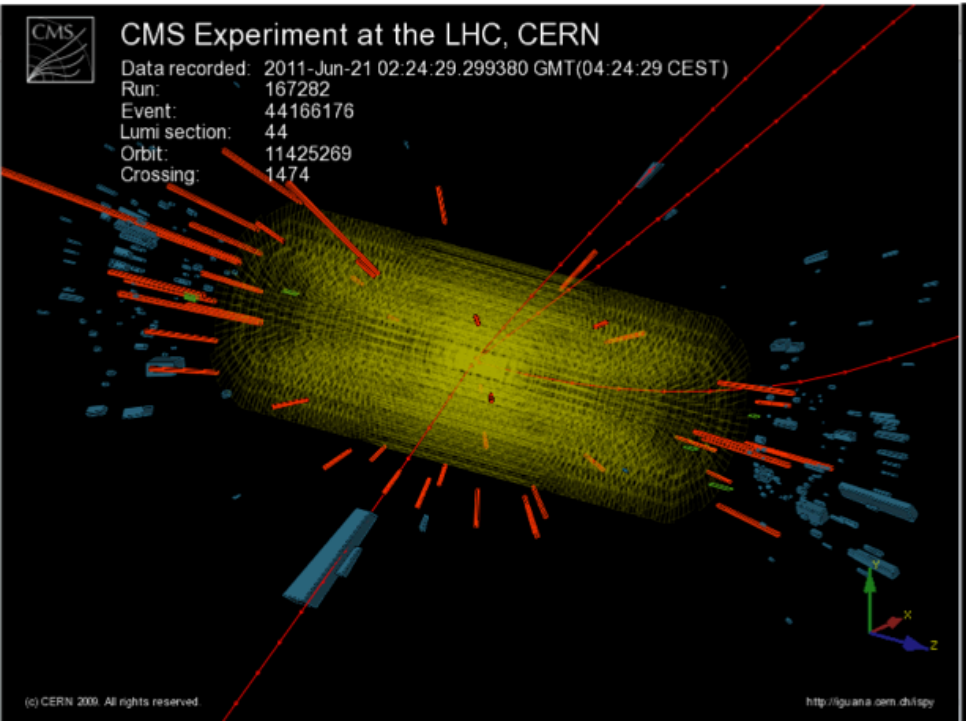


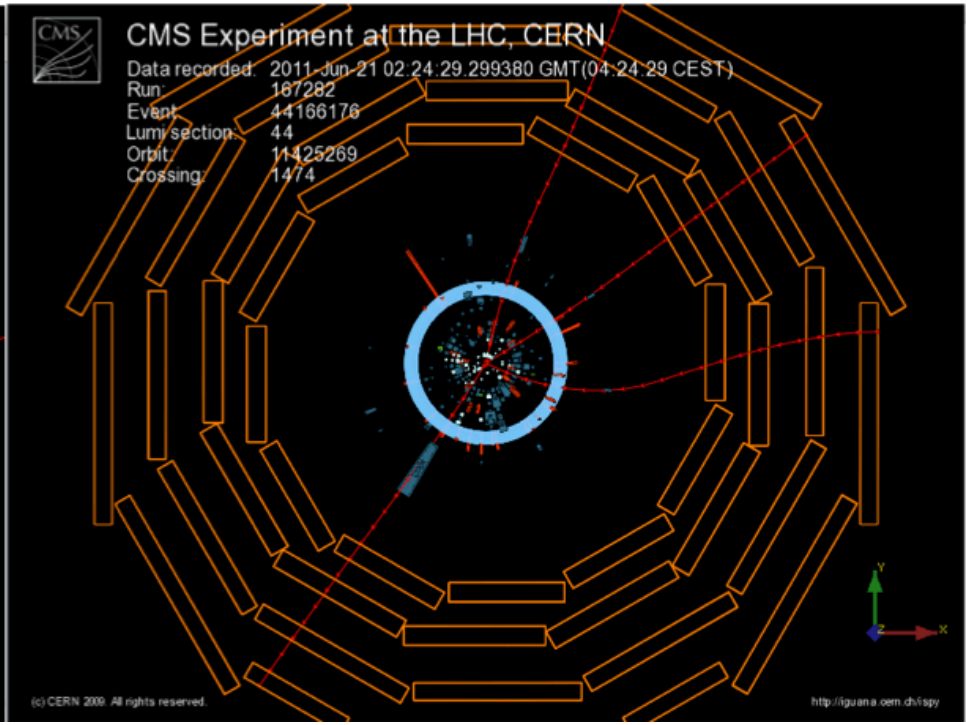
Local p-value



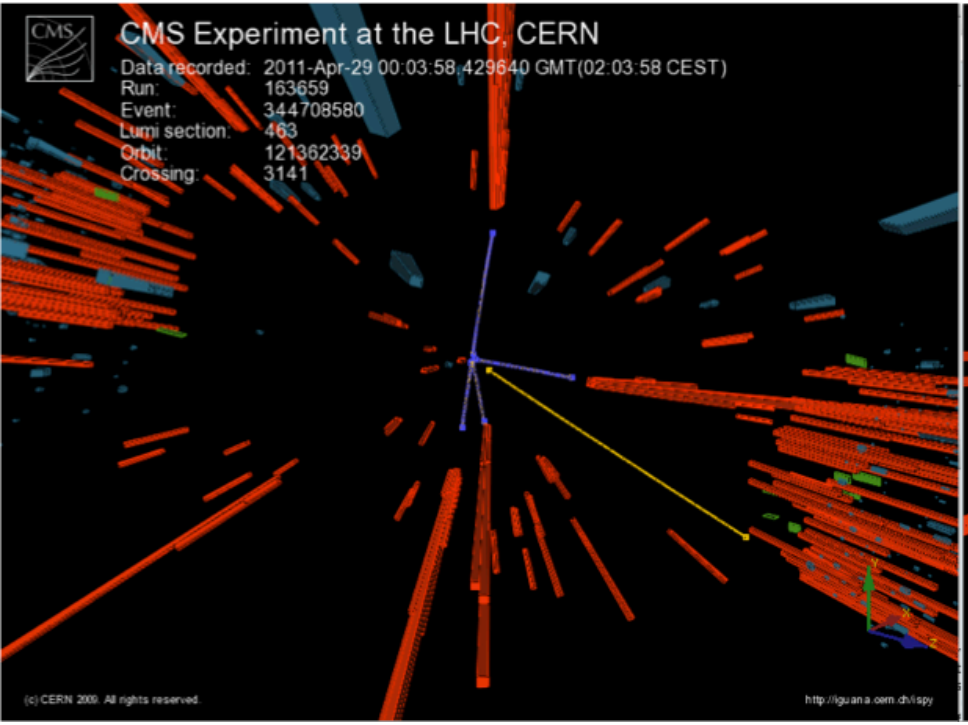


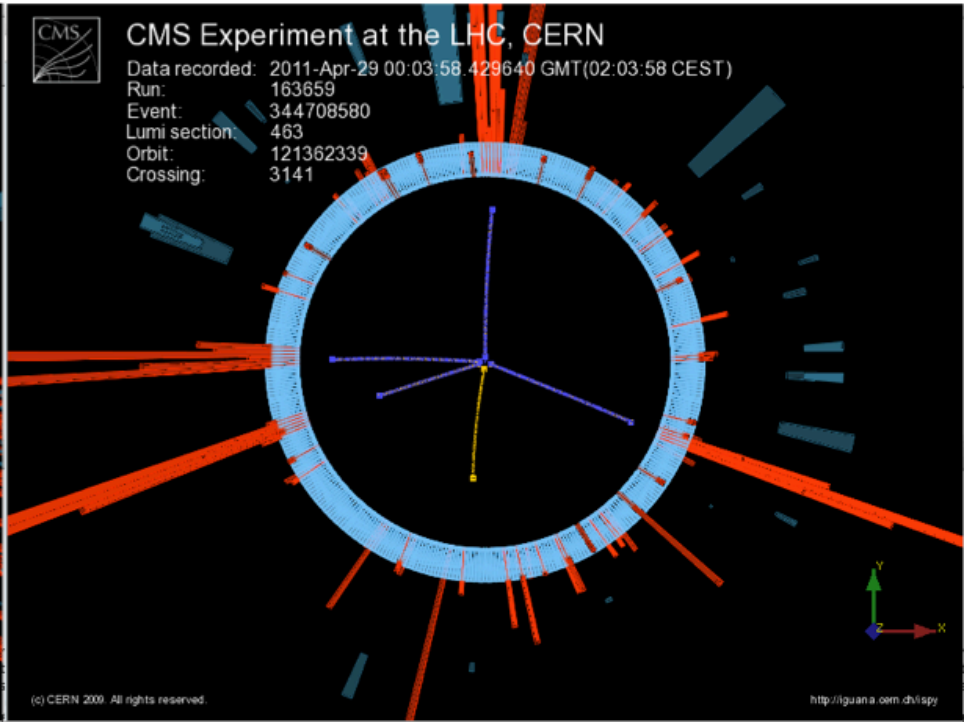
Event display – 4mu 7 TeV



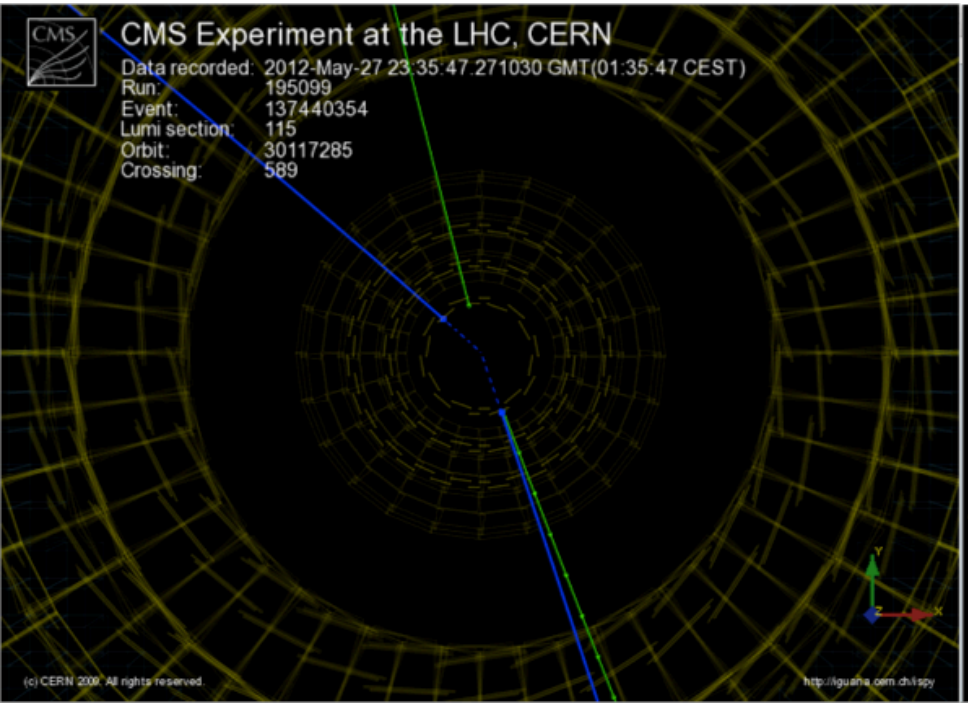


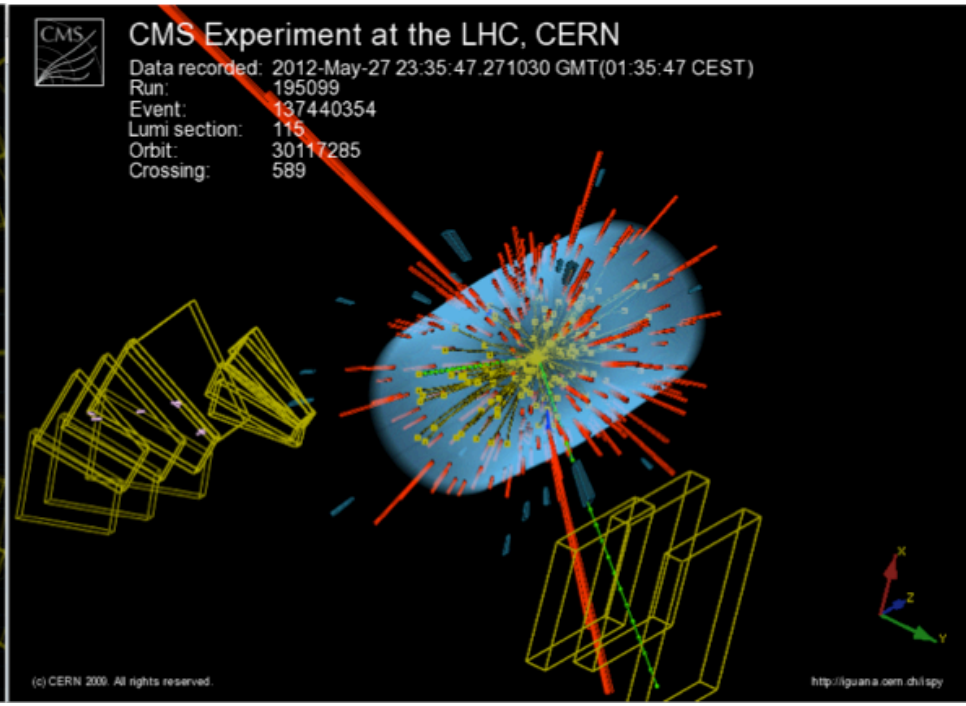
Event display – 4e 7 TeV





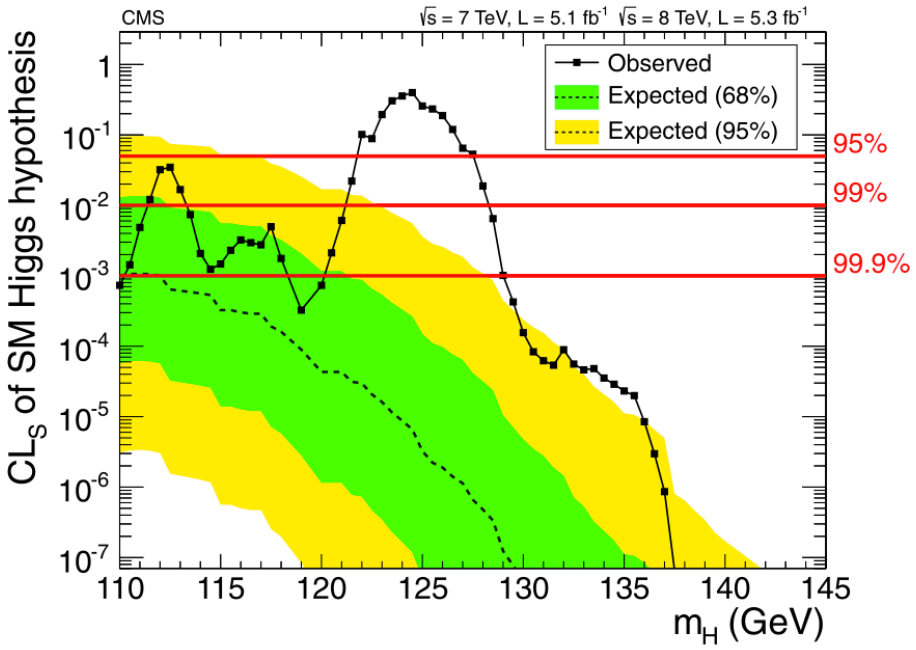
Event display – 2e2mu 8 TeV

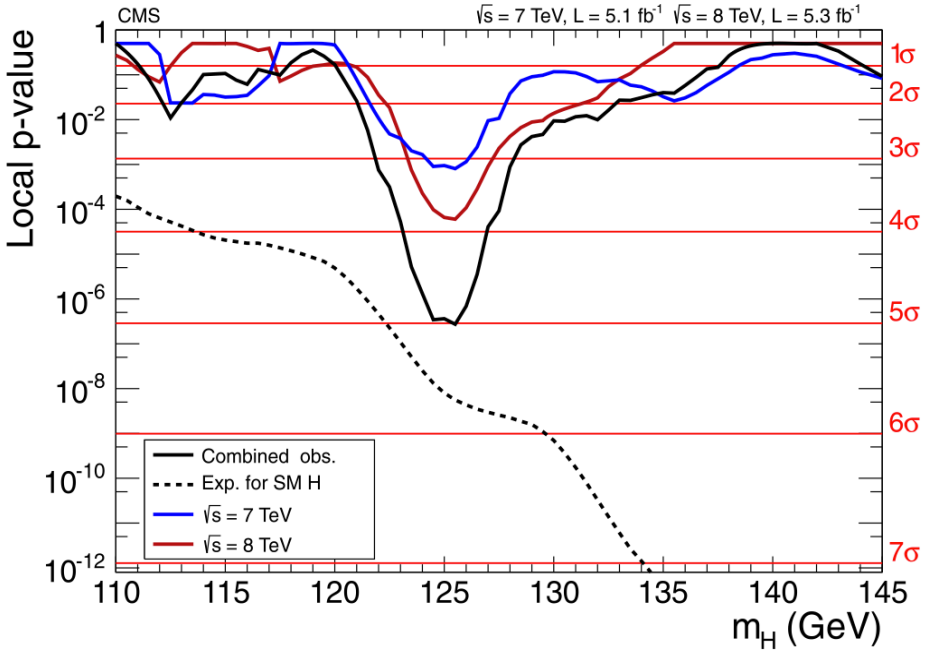




2012: a historic year

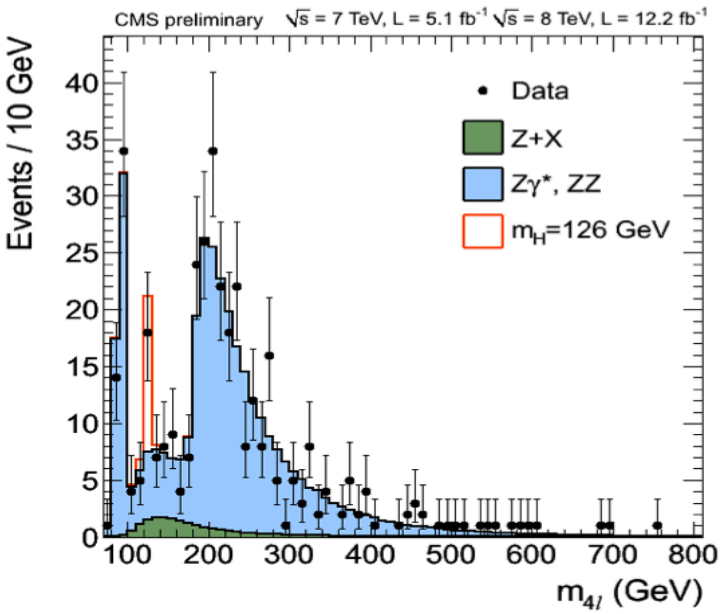
- 4th July 2012: the CMS and ATLAS collaborations announce the observation of a **new boson** with a **mass ~125 GeV**, with a local **significance of ~5σ** in both experiments

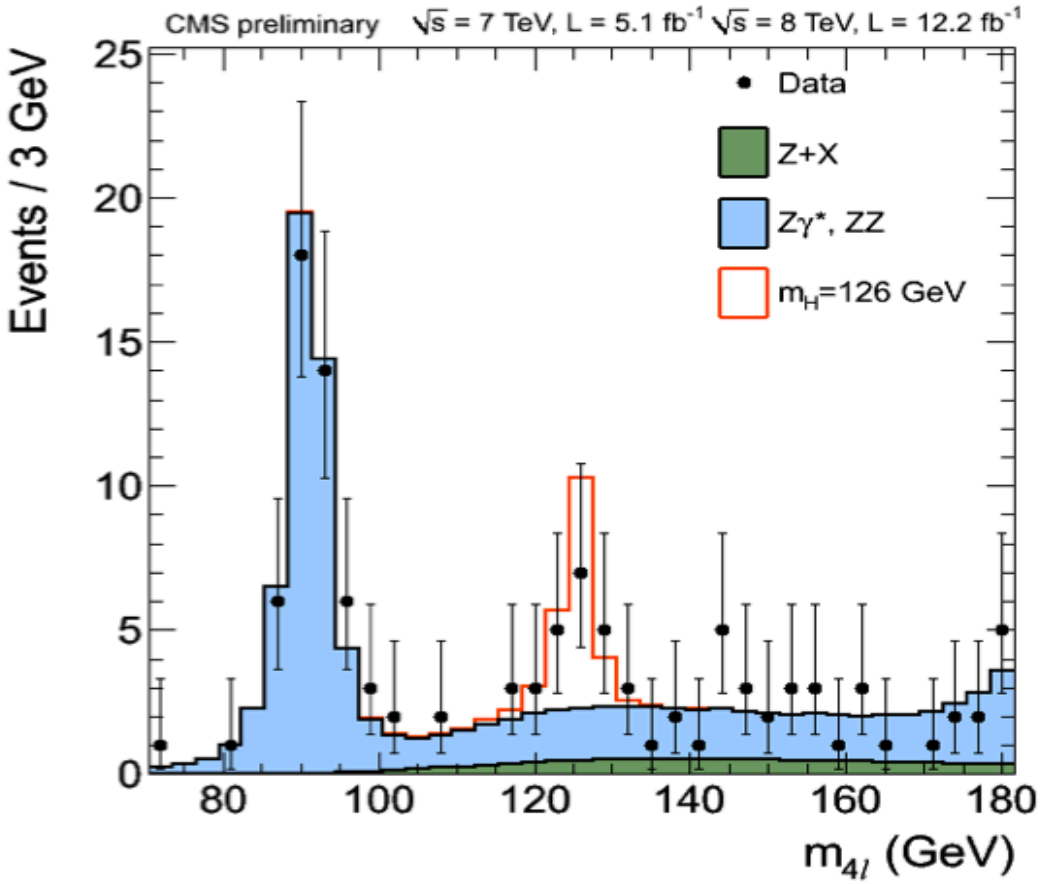




Official CMS results on H→ZZ→4l

- The previous results must be considered a cross-check of the official ones





[121.5,130.5]	7 TeV exp	8TeV exp	obs
4e	1.25	2.20	3
4μ	2.09	4.26	6
2e2μ	3.14	5.97	8
Total	**6.48**	**12.43**	**17**

Official CMS results on H→ZZ→4l

- Largest excess wrt backround at m_H = 126 GeV: **local significance 4.5σ**
- Excluded mass range at 95% CL: **[113,116] U [129,720] GeV**
- **cross section**: $\mu = 0.8^{+0.35}_{-0.28}$
- **mass**: $m_H = 126.2 \pm 0.6\ (stat) \pm 0.2\ (syst)$